MEDICAL ROBOTS

Autonomous robotic intraocular surgery for targeted retinal injections

Gui-Bin Bian[1]*†, Yawen Deng[2]†, Zhen Li[1], Qiang Ye[1], Yupeng Zhai[3], Yong Huang[4], Yingxiong Xie[4], Weihong Yu[5], Zhangwanyu Wei[5], Zhangguo Yu[2]

Intraocular surgery is challenged by restricted environmental perception and difficulties in instrument depth estimation. The advent of autonomous intraocular surgery represents a milestone in medical technology, given that it can enhance surgical consistency that improves patient safety, shorten surgeon training periods so that more patients can undergo surgery, reduce dependency on human resources, and enable surgeries in remote or extreme environments. In this study, an autonomous robotic system for intraocular surgery (ARISE) was developed, achieving targeted retinal injections throughout the intraocular space. The robotic system achieves intelligent perception and macro/microprecision positioning of the instrument throughout the intraocular space through two key innovations. The first is a multiview spatial fusion that reconciles imaging feature disparities and corrects dynamic spatial misalignments. The second is a criterion-weighted fusion of multisensor data that mitigates inconsistencies in detection range, error magnitude, and sampling frequency. Subretinal and vascular injections were performed on eyeball phantoms, ex vivo porcine eyeballs, and in vivo animal eyeballs. In ex vivo porcine eyeballs, 100% success was achieved for subretinal ($n = 20$), central retinal vein (CRV) ($n = 20$), and branch retinal vein (BRV) ($n = 20$) injections; in in vivo animal eyeballs, 100% success was achieved for subretinal ($n = 16$), CRV ($n = 16$), and BRV ($n = 16$) injections. Compared with manual and teleoperated robotic surgeries, positioning errors were reduced by 79.87 and 54.61%, respectively. These results demonstrate the clinical feasibility of an autonomous intraocular microsurgical robot and its ability to enhance injection precision, safety, and consistency.

INTRODUCTION

The human eyes are the primary sensory organs, responsible for processing more than 80% of external information (*1*–*3*). More than 2.2 billion people worldwide have eye diseases (*4*, *5*), with vitreoretinal conditions being a leading cause of blindness (*6*). Because of the intricate intraocular structure and the extremely limited operating space, vitreoretinal surgery requires delicate surgical instruments and high precision (*7*), placing substantial technical burden on the surgeon (*8*). Advancements in subretinal gene therapy have made more ocular diseases treatable. However, techniques like subretinal injection and venous cannulation require highly precise localization and microscale manipulation because of the retinal pigment epithelium's 25-μm thickness and vessels' typical diameter of less than 150 μm (*9*). Because natural tremors, muscle twitches, and low-frequency drifts in human hands cause deviations greater than 100 μm, the precision required in surgery surpasses the physiological limits of manual control (*10*). Robot-assisted surgery is an emerging technique that enhances surgical safety and reduces intraoperative complications by improving positioning accuracy through motion scaling and tremor suppression while also enabling instruments to remain steady and immobilized for prolonged periods (*11*, *12*). It also demonstrates distinct advantages in ophthalmic surgery (*13*). Typical ophthalmic surgical robotic systems comprise platforms such as the KU Leuven comanipulator system (*14*–*16*), PRECEYES (*17*, *18*), iRAM!S (*19*), SHER (*20*–*23*), IRISS (*24*–*26*), and RVRMS (*27*). Many of these platforms offer multi–degree-of-freedom manipulation for precise control in constrained intraocular spaces. In particular, head-mounted surgical robots (*10*) can effectively mitigate head-induced relative motion during subretinal injections. Current research primarily emphasizes technological advancements in areas including system flexibility (*28*), tremor suppression (*29*), microforce sensing (*30*, *31*), and motion control accuracy (*32*) in ophthalmic surgical robots. Most research relies on simulated silicone eyeballs and ex vivo animal models, with fewer live animal studies. Although most ophthalmic robotic systems are still in the preclinical phase, both PRECEYES and the KU Leuven comanipulator system have advanced to early human clinical evaluation, demonstrating feasibility in procedures such as subretinal injection and retinal vein cannulation (*15*, *17*, *18*).

Autonomy in surgical robotics, defined as the system's ability to perceive, plan, and execute tasks with minimal human intervention, is increasingly regarded as a critical advancement in microsurgical disciplines such as ophthalmology (*12*, *33*–*35*). In vitreoretinal procedures, surgical robots must achieve micrometer-level precision, stable tool-tissue contact, and reliable intraocular navigation (Fig. 1). Although multi–degree-of-freedom teleoperated robotic systems reduce tremor compared with manual surgery, their capability for manipulation remains complex. Surgeons must simultaneously coordinate instrument position and orientation adjustments in the three-dimensional (3D) space, which places considerable demands on spatial reasoning. Meanwhile, teleoperated robotic systems are challenged by a steep learning curve that necessitates extensive surgeon training and undermines consistency across operators and surgical outcomes (*11*). Autonomous systems aim to overcome these limitations by enabling closed-loop perception, planning, and control, with the ultimate goal of enabling fully autonomous execution of microsurgical tasks.

[1]State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. [2]School of Mechatronic Engineering, Beijing Institute of Technology, Beijing, China. [3]College of Integrated Circuits, Taiyuan University of Technology, Taiyuan, China. [4]School of Optics and Photonics, Beijing Institute of Technology, Beijing, China. [5]Department of Ophthalmology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China.
*Corresponding author. Email: guibin.bian@ia.ac.cn
†These authors contributed equally to this work.

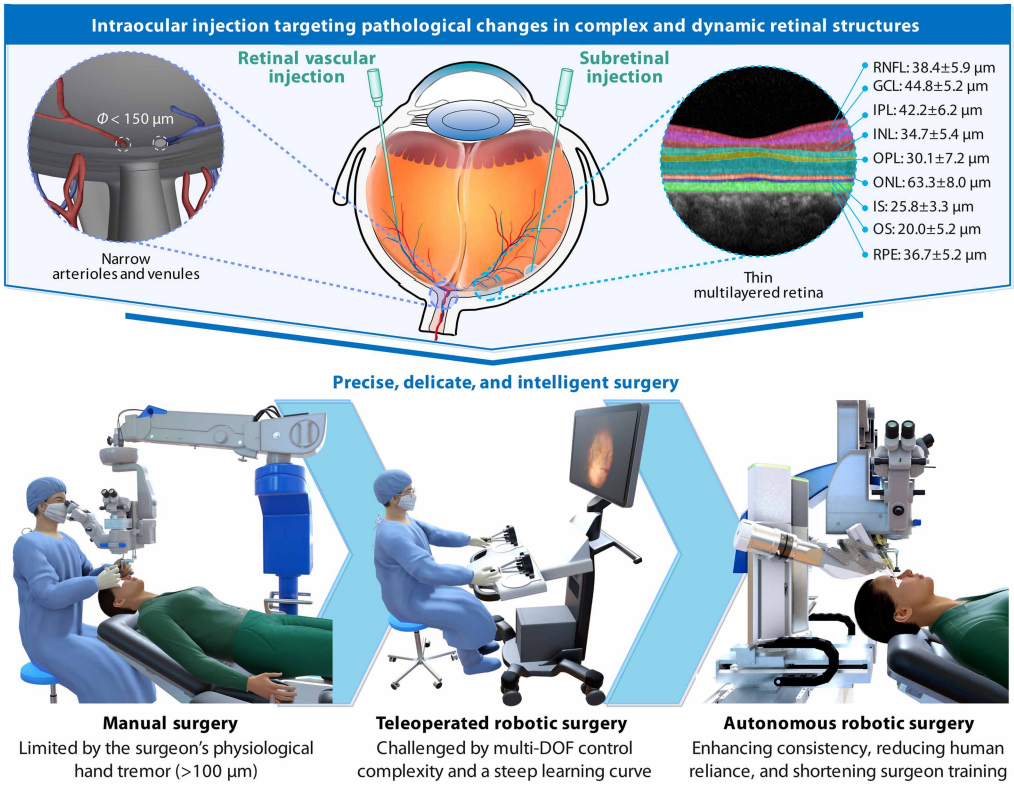


Fig. 1. Intraocular structures and evolution of operative techniques in intraocular surgery. The intricate intraocular structures demand precise, delicate, and intelligent surgical approaches. With the advancement of treatments, manual and teleoperated robotic surgeries can no longer fully meet clinical requirements. Autonomous robotic surgery enhances surgical consistency, improves patient safety, reduces reliance on human resources, and minimizes human errors, thereby supporting safer and more reliable treatments. RNFL, retinal nerve fiber layer; GCL, ganglion cell layer; IPL, inner plexiform layer; INL, inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; IS, inner segment; OS, outer segment; RPE, retinal pigment epithelium; DOF, degree of freedom.

Achieving autonomy in delicate intraocular environments requires reliable decision-making and precise actions, offering notable benefits despite the challenges. These include enhanced consistency that improves patient safety and outcomes, accelerated training that enables more patients to receive microsurgical care, and reduced operator variability that optimizes resource allocation in settings constrained by workforce shortages or geographic limitations (*36*). Existing autonomous surgical robotic systems include the clinically implemented robotic milling system for total knee arthroplasty (*37*), whereas systems such as the intestinal anastomosis system for laparoscopic surgery (*38*–*40*), the intracardiac catheter navigation system in cardiac procedures (*41*, *42*), and the needle guidance system for pulmonary interventions (*43*) are still under investigation in animal studies. In addition, the SRT-H system (*44*) demonstrated step-level autonomy in ex vivo cholecystectomy through a hierarchical framework integrating language- conditioned planning and motion control. In contrast, ophthalmic robotic systems remain largely surgeon operated and exhibit only limited autonomy. Full automation in vitreoretinal surgery has not yet been achieved, with current research focusing on specific subtasks such as vessel puncture (*21*, *25*), illumination guidance (*22*), or semiautomated actions (*23*). For example, Zhang *et al.* (*21*) integrated the SHER robotic system with intraoperative optical coherence tomography (iOCT) for autonomous needle navigation, automating only the puncture step while keeping positioning and injection manual. Full autonomy in ophthalmic surgery remains challenging because of limited perception of the intraocular environment and insufficient depth awareness, both of which hinder precise instrument localization.

Ophthalmic surgical perception primarily relies on surgical microscopes, which conventionally serve as the main visualization modality during microsurgery but are constrained by a limited field of view (FOV) and the inability to reveal microstructures and tissue planes beneath the retinal surface (*12*). Even with wide-angle fundus lenses, the FOV of the posterior segment is restricted to ~50°. Microscope-based ophthalmic surgery is impaired by shadows, reflections, blurred contours, occlusion, uneven lighting, and lens artifacts such as particles and droplets (*45*, *46*). OCT provides high-resolution cross-sectional retinal images (*47*), and iOCT has been introduced into ophthalmic operating rooms (*48*) as a visualization aid (*26*). Building on this, prior studies have explored iOCT-guided robotic approaches for subretinal injection, including autonomous navigation (*49*), tool localization with model predictive control (*21*), and human-in-the-loop visualization with uncertainty overlay to enhance surgeon supervision (*50*). Nevertheless, the FOV of traditional OCT is limited (~4 mm by 4 mm) (*51*), which restricts imaging to localized regions.

Current methods for localizing instrument depth during ophthalmic surgery remain limited in scope and fail to fully account for the complexity of real surgical environments. Common approaches include robot kinematics, instrument projection, and OCT-based axial imaging. Kinematics infers depth from joint positions and the scleral trocar (*32*) but requires manual retinal plane estimation and

achieves only 100-μm accuracy. Projection methods estimate relative depth via geometric relations (*52*–*54*), constrained by microscope FOV and instrument proximity, yielding only relative pixel-based depth. Shadow-target depth is often fitted with planes or circles, neglecting retinal irregularities. OCT provides cross-sectional images along the instrument axis (*55*) but is limited by the narrow axial FOV and lack of lateral integration.

To address the limitations mentioned above, this study has three main contributions. First, an autonomous robotic system for intraocular surgery (ARISE) has been developed, enabling fully autonomous operation within the intraocular space. The system performs subretinal and intraretinal injections, validated on eyeball phantoms, ex vivo porcine eyeballs, and in vivo animal eyeballs (Movie 1). In ex vivo porcine eyeballs, 100% success was achieved for subretinal ($n = 20$), central retinal vein (CRV) ($n = 20$), and branch retinal vein (BRV) ($n = 20$) injections; in in vivo animal eyeballs, 100% success was also achieved for subretinal ($n = 16$), CRV ($n = 16$), and BRV ($n = 16$) injections. ARISE also demonstrated substantially improved positioning accuracy, reducing errors by 79.87 and 54.61% compared with manual and teleoperated approaches, respectively.

Second, a multiview spatial fusion is proposed to overcome imaging heterogeneity and dynamic spatial misalignment in multimodal intraocular imaging. This method addresses both challenges through an interactive spatial alignment module and a cross-modal shift prediction module. By integrating information from both modules, ARISE constructs a dynamic 3D intraocular map. This map substantially reduces sensitivity to variations in microscope image quality and effectively compensates for regions that are undetectable by existing single-modal approaches.

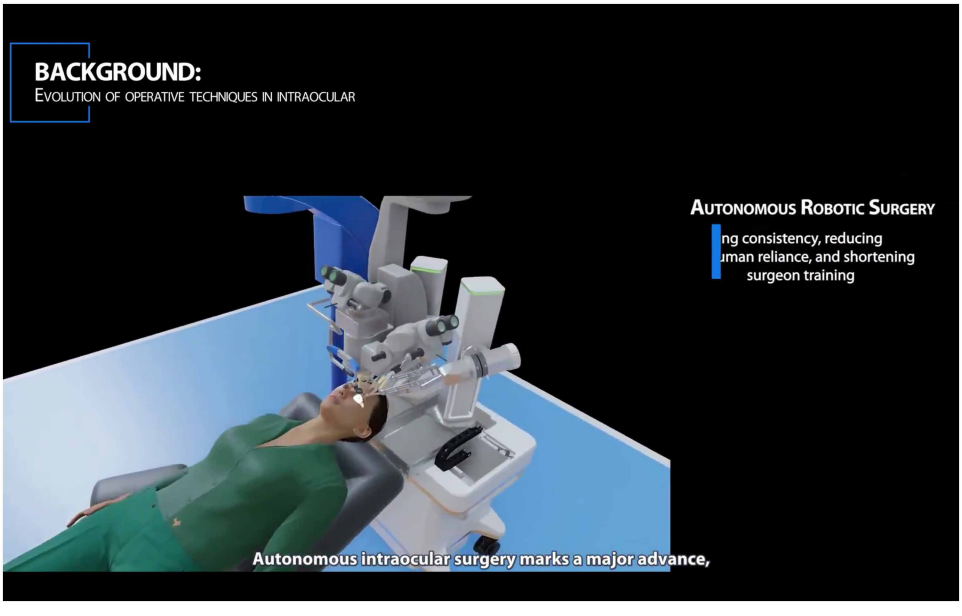
Third, a criterion-weighted fusion of multisensor data is proposed to address differences in detection range, error magnitude, and sampling frequency. This method achieves precise localization of the instrument tip at both macro- and microlevels, thereby enabling precise interaction and real-time adjustment within the complex intraocular environment. It comprises multisource data normalization and multicriterion evidence source fusion. Multisource data normalization aligns the outputs of the macro- and microlevel localization algorithms through cotarget registration and multifrequency interpolation, whereas fusion of these localization results applies the best-worst method across four criteria to derive weights and compute the final score.

RESULTS

ARISE

ARISE is composed of a high-precision dual-manipulator microsurgical robot, 3D intraoperative navigation, and a human-observed force-position-imaging-hybrid safety control module. Figure 2 shows the overall surgical scene, including visible components of ARISE, whereas movie S1 provides a comprehensive overview of its architecture and operational workflow. ARISE could perform intraocular surgeries such as subretinal and intraretinal vascular injections.

Inspired by human bimanual operation, the dexterous dual-manipulator microsurgical robot was designed for intraocular procedures (Fig. 3A). The right manipulator holds an injection needle equipped with force and distance sensing, whereas the left manipulator holds a medical lighting fiber for intraocular illumination. This dual-manipulator configuration enabled a compact surgical field and reduced spatial calibration error through shared base coordinates, thereby enhancing accuracy in intraocular procedures. The robotic platform consists of an *XYZ* three-axis motion module and a fixed-point mechanism based on a dual-parallel four-bar linkage with a distal mechanical constraint, offering 12 degrees of freedom. Its workspace (298.4 mm by 204.4 mm by 176.0 mm) encompasses the full operational area surrounding the human head. This provides sufficient clearance to prevent collisions with surgical devices such as microscopes and vitrectomy instruments. The end effector defined the remote center of motion (RCM) at the scleral trocar insertion site, achieving a repeatability of 7 μm (measured by a laser tracker, AT960-LR). It maintained precise spatial movement under a 2-kg tip load at speeds up to 40 mm/s, enabling minimally invasive, accurate operations across the posterior segment of the eye. The injection needle on the right manipulator was constructed from a superelastic nickel-titanium tube. It integrated dual–fiber Bragg grating (FBG) optical fibers and an iOCT A-scan probe. These components enabled simultaneous sensing of puncture force and distance to the target region. The rear end was connected to a sealed, antidisturbance drug delivery conduit to ensure stability during fluid injection. The outer diameter of the puncture cannula was 60 μm. The motion control resolution at the needle tip reached 0.5 μm, with a force sensitivity of 2 mN. The iOCT A-scan axial FOV at the tip was 3.62 mm, with a resolution of 5.3 μm. The iOCT module comprised a B-scan integrated into the microscope's optical path and an A-scan embedded within the injection needle. The B-scan provides wide-field cross-sectional views of the needle tip and retinal structures, whereas the A-scan offered



Movie 1. ARISE achieves autonomous targeted subretinal and retinal vascular injections and is validated on eyeball phantoms, ex vivo porcine eyeballs, and in vivo animal eyeballs.

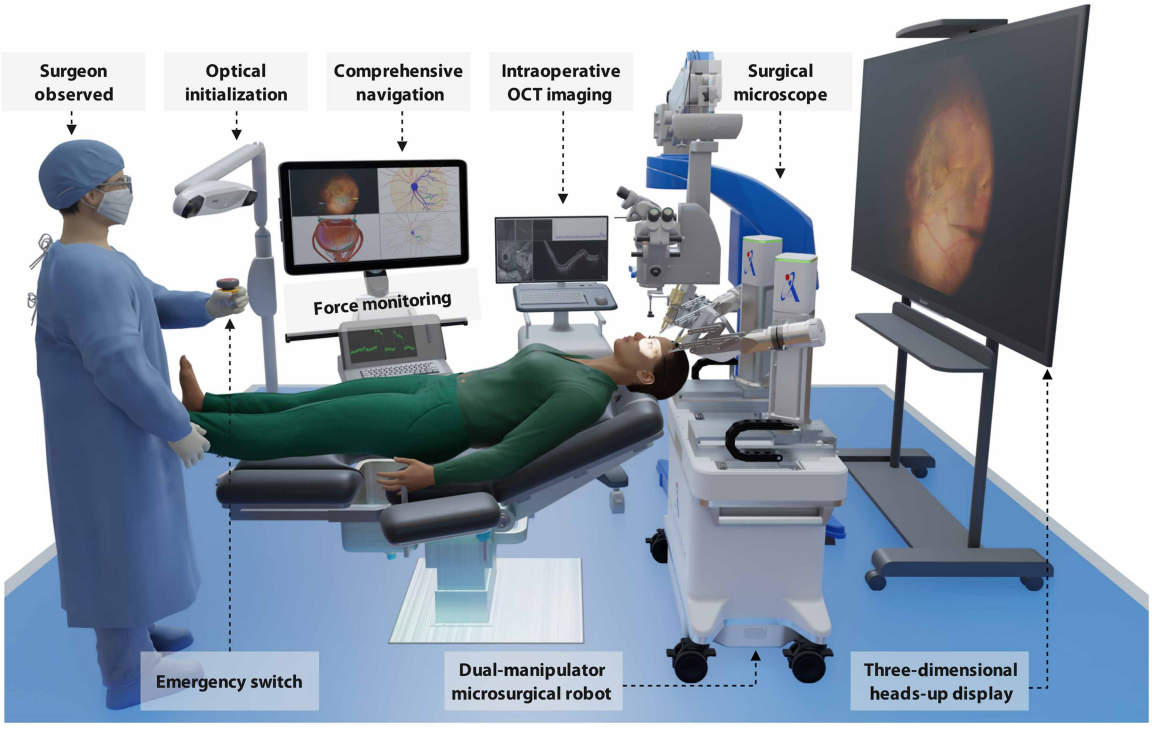


Fig. 2. Autonomous robotic system for intraocular surgery (ARISE). A high-precision, dexterous dual-manipulator microsurgical robot performs autonomous retinal injection guided by intraoperative 3D navigation, with the procedure conducted under human-observed force-position-imaging-hybrid safety control.

localized, micrometer-scale depth information relative to surrounding tissues. These capabilities supported safe and precise delivery of drugs into intravitreal and subretinal spaces.

A 3D intraoperative navigation module integrating multimodal information was developed to support dynamic global 3D map construction (Fig. 3B), macro/micro-intraocular instrument-tip localization (Fig. 3C), and multiconstraint path planning (Fig. 3D). This navigation module autonomously positioned the injection needle of ARISE at the surgeon-defined target, whereas the lighting fiber adjusted its orientation to maintain optimal visibility. The navigation integrated diverse sensor data, including preoperative OCT and fundus color images, actuator states from ARISE, iOCT A-scan and B-scan, intraoperative microscope views, and optical tracking of external anatomical landmarks or instruments.

Human-observed force-position-imaging-hybrid safety control supported dynamic sensing and evaluation of instrument-tissue interaction, enabling real-time risk assessment and early warning to enhance surgical safety with ARISE (Fig. 3E). A domain knowledge–guided, multitask instrument state perception approach integrated the robot's global pose, intraoperative relative pose estimation, and target positioning. This integration was used to identify surgical steps and to continuously monitor, predict, and constrain the instrument's trajectory and speed throughout the procedure. Low-latency, high-sensitivity force sensing enabled real-time detection of membrane and vessel puncture, along with precise feedback and puncture depth control. Human-robot collaborative monitoring was enhanced by real-time visual feedback and dynamic information fusion. This mechanism provided immediate warning and supported seamless transitions between autonomous and manual operation, thereby improving intraoperative safety and control.

Multiview spatial fusion enabled the construction of an intraoperative dynamic global 3D map (Fig. 3B), supporting intraoperative environmental perception during ARISE procedures. First, PRTransUNet, previously proposed by our research group (*56*), was applied to segment features from preoperative and intraoperative multimodal images. These extracted features, along with their confidence weights, were processed by two coordinated modules: an interactive spatial alignment module and a cross-modal shift prediction module. The former combined a local-global spatial transformer and an authenticity predictor to construct lateral-to-axial FOV mapping, resolving modality disparities and enabling the spatiotemporal alignment of multisource intraocular imaging. The latter used an instance perception adapter to track various local displacements and subsequently predicted global shifts on the basis of geometric relationships and a matching matrix. The local-global spatial transformer integrated three algorithms. The first algorithm was a coarse intraocular reconstruction based on pupil-trocar geometry. The second algorithm was the registration of global microfeatures in fundus images. The third algorithm was a lateral field feature mapping–based 3D fundus image registration. Together, these algorithms initialized retinal structure modeling, recovered missing details, refined intraoperative retinal reconstruction, and provided stable landmarks for ARISE navigation. The authenticity predictor assigned confidence weights to each unimodal feature in the fusion region across three algorithms, generating more discriminative multimodal representations. The fusion feature set $\mathbf{F}_{\text{fusion}}$ represented the integrated multimodal features and was computed as follows

$$\mathbf{F}_{\text{fusion}} = \{\mathbf{F}_k | k \in \{1, 2, \ldots, n\}\} \quad (1)$$

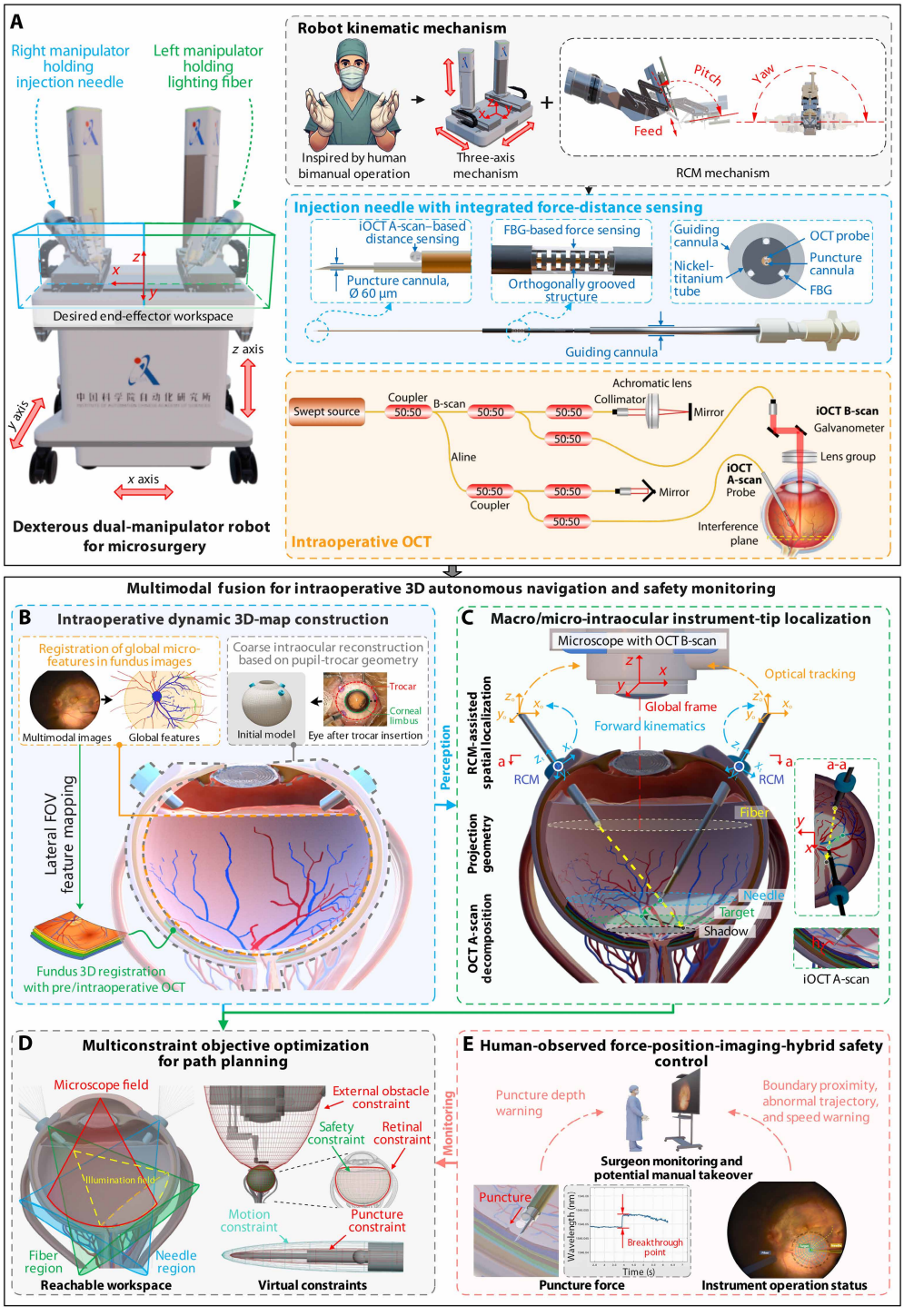


Fig. 3. Components of ARISE. (**A**) The microsurgical robot was designed with a dual-manipulator configuration offering 12 degrees of freedom, inspired by human bimanual operation. The left manipulator held a lighting fiber, whereas the right manipulator held a custom-developed injection needle whose tip diameter was only 60 μm and incorporated OCT A-scan–based distance sensing and microforce sensing. (**B**) By integrating preoperative and intraoperative data from both lateral and axial FOVs, a personalized dynamic global 3D map was constructed intraoperatively for each eye through multiview spatial fusion. (**C**) With the global map, criterion-weighted fusion of multisensor data enabled layered instrument pose perception from the anterior segment to the retinal fundus, along with interaction modeling of the intraocular environment. (**D**) By incorporating constraints from the RCM, microscope FOV, illumination geometry, tissue protection margins, and instrument motion limits, multiple constraints delineated obstacle regions and safe operation zones, forming a global configuration space to support path planning. Ellipsoidal modeling reduced computational complexity and facilitated safe, precise injection trajectory generation. (**E**) The human-observed force-position-imaging-hybrid safety control module dynamically assessed instrument-tissue interaction, monitored puncture events in real time, limited puncture depth, and supported seamless transition between robotic and manual control.

where each fused feature $\mathbf{F}_k$ was calculated by combining features from individual modalities, weighted by modality confidence scores $w_{k,m}$ and authenticity confidence scores $\boldsymbol{C}_{\text{auth},k,m}$

$$\mathbf{F}_k = \sum_{m=1}^{M} w_{k,m} C_{\text{auth},k,m} \mathbf{F}_{k,m} \tag{2}$$

$\mathbf{F}_{k,m}$ denotes the kth feature to be extracted, which includes fundus vessels, retinal layers, surgical instruments, and other features. $\boldsymbol{C}_{\text{auth},k,m}$ was defined by the similarity between intraoperative microscopic images and preoperative fundus photographs mapped to the global space. It evaluated the authenticity of intraoperative features, given that their confidence weights had not yet been validated in practical applications. The modality confidence weight $w_{k,m}$ quantified the relative reliability of each modality's extracted features. It was calculated via a softmax normalization strategy to dynamically prioritize features from higher-quality imaging sources. This approach directly mitigated issues arising from inconsistent image quality or partial occlusions during surgery.

$$w_{k,m} = \frac{\exp\left(C_{k,m}\right)}{\sum_{m'=1}^{M} \exp\left(C_{k,m'}\right)} \tag{3}$$

where M is the total number of modalities and $\boldsymbol{C}_{k,m}$ is the confidence score of the kth feature in the mth modality, derived from segmentation outputs. The cross-modal shift prediction module tracked the planned injection point, incorporated outputs from the local-global spatial transformer, and corrected spatial misalignments from tissue displacement, microscope movement, and instrument manipulation. The dynamic map updated at 14 Hz within a localized retinal area of 12 mm by 9 mm by 9 mm and required ~44 s for construction and initialization of the full-eye 3D map. The constructed dynamic global 3D map achieved a lateral resolution of 4.7 μm within the microscope's FOV and 20 μm outside the field. Its depth resolution was 6 μm, achieving 99.08% accuracy in replicating true retinal depth.

Criterion-weighted fusion of multisensor data enabled precise intraocular instrument localization during surgery (Fig. 3C). It systematically integrated intraoperative data from multiple sensors. To address differences in detection range and sampling frequency, a dedicated normalization module performed resampling and spline interpolation to ensure temporal and spatial consistency. The positioning strategy used three cascaded and complementary levels, each aligned with a specific anatomical region, to achieve robust localization across the entire surgical workspace. Outside the microscope's FOV, the extrinsic spatial localization algorithm assisted by the RCM fused optical tracking with robotic kinematics for position estimation. In the posterior segment, the end-effector localization algorithm segmented contour and shadow features to model the instrument and recovered 3D coordinates with sub–10-μm accuracy. Near the retina, the relative localization algorithm based on iOCT analyzed A-scan data to reconstruct the needle tip's 3D position with micrometer-level precision. Localization results from the three levels were integrated in a unified coordinate space to ensure accurate and robust real-time tracking throughout the procedure. To evaluate the reliability of each sensing modality under varying conditions, the fusion module computed a weighted evidence score

$$S = w_{\text{R}} \cdot R + w_{\text{E}} \cdot E + w_{\text{T}} \cdot T + w_{\text{C}} \cdot C \tag{4}$$

Each component reflected a measurable aspect of sensor performance. Reliability ($\boldsymbol{R}$) was derived from feature confidence, such as segmentation certainty or tracking stability. Effectiveness ($\boldsymbol{E}$) was evaluated on the basis of the empirical error distribution of each sensor. Real-time performance ($\boldsymbol{T}$) corresponds to sampling frequency and latency. Consistency ($\boldsymbol{C}$) was quantified using the Mahalanobis distance across modalities. The optimal weight vector $\mathbf{w} = \left[w_{\text{R}}, w_{\text{E}}, w_{\text{T}}, w_{\text{C}}\right]^T$ was determined using the best-worst method. This approach minimized subjective bias by comparing the most and least important criteria against all others, thereby constructing a consistent pairwise preference matrix $\mathbf{A}$ and solving the eigenvalue-based optimization.

$$\mathbf{A} \cdot \mathbf{w} = \lambda_{\max} \cdot \mathbf{w} \tag{5}$$

This formulation ensured that each fusion component contributed in a quantitatively validated and context-specific manner, enabling robust, real-time, and anatomically aware end-effector localization across dynamic intraoperative scenes.

Multiconstraint objective optimization enabled the planning of illumination and injection paths to support precise, safe, and efficient subretinal and intraretinal vascular injections (Fig. 3D). The approach accounted for constraints including the RCM, microscope FOV, optimal illumination zone, intraocular tissue protection boundaries, and external instrument motion limitations. It defined obstacle and safe regions on the basis of multiple constraint conditions, constructing a global configuration space that links the RCM, instrument operating range, and injection targets. Ellipsoidal modeling approximated obstacle boundaries and instrument tips to reduce computational load and dynamically adapted to varying precision requirements across different surgical phases. Coordinated motion of the injection needle and lighting fiber was maintained to ensure sustained optimal illumination during task execution. Path planning was based on the rapidly exploring random tree algorithm and adapted to the constraint framework for path validation and optimization. This strategy achieved a comprehensive balance of safety, cooperation, and efficiency throughout the planned trajectory.

Autonomous procedures for targeted retinal injection

ARISE was designed to perform subretinal and intraretinal vascular injections at targets predefined by the surgeon (Fig. 4) through three stages: preoperative preparation, target positioning, and puncture and injection. The latter two stages were fully autonomous and were executed by ARISE.

During preoperative preparation, a retinal map was generated as the surgeon teleoperated the dual-manipulator robot to traverse the scleral trocar, insert the lighting fiber and needle, and select the injection target. The z axis aligned with the microscope's optical axis, whereas the x and y axes lay in the orthogonal plane, defining the global coordinate frame. The optical tracker recorded the microscope's FOV center and instrument pose at the trocar entry to locate the RCM in global coordinates. Combining the pupil-trocar geometry with ocular anatomy yielded an initial global map. Preoperative fundus photographs and OCT images were also acquired to refine the global map.

The injection target positioning stage was structured into three substages. In stage 2.1, the lighting fiber and injection needle were positioned via an extrinsic spatial localization algorithm assisted by the RCM, guided into the microscope's FOV, and adjusted to illuminate the target region. In stage 2.2, the robot's motion was calibrated

by executing an asymmetric quadrilateral trajectory that captured kinematic characteristics under varied motion. The robot followed each edge with uniform motion, pausing at corners to record data. During this process, the system collected robot joint positions, pixel coordinates of the visible guiding cannula tip under the microscope (as the needle tip remained retracted), and the instrument pose from optical tracking. The data enabled error quantification, which optimized the kinematic model and refined navigation by correcting joint offsets and aligning imagery, tracking, and motion. By matching preoperative and intraoperative lateral FOVs with iOCT, the global map was refined using fused 3D retinal data to achieve high-precision surgical localization. In stage 2.3, the needle descended toward the target, and shadow detection triggered the end-effector localization algorithm, which guided the needle into the iOCT field. When it was within a range of 1 to 2 mm from the retina, the needle extended from the cannula, and the iOCT-based relative localization algorithm performed fine adjustment to align the tip with the target.

During the puncture and injection, the needle tip followed a navigation-guided path and was continuously adjusted in real time to puncture the retina and deliver agents into subretinal or intraretinal targets. The safety control module adjusted force thresholds according to the biomechanical properties of the target, such as the internal limiting membrane or vessel wall. This enabled detection of force peaks signaling membrane penetration. Upon puncture detection, ARISE estimated tissue rebound. It compared insertion depth at force release with the thickness of retinal or vascular layers from intraoperative OCT B-scans. This quantified tissue compression and recovery. Real-time constraints prevented overpenetration and ensured safe tissue traversal. Meanwhile, the dynamic global map monitored intraocular motion. It triggered alerts for sudden displacement spikes, often from respiratory movements during anesthesia recovery. In such cases, the surgeon could take control or withdraw the instrument. Retraction (less than 3 s) was defined as the time between receipt of the command, during which the instrument remained in contact with the retina, and its complete withdrawal from the intraocular space, as shown in movie S1. This ensured safe disengagement during unexpected intraoperative motion.

Autonomous injection localization in eyeball phantoms

A silicone eyeball model, simulating the complex retinal vasculature of the human eye, was developed to evaluate the autonomous positioning capability of ARISE in a realistic intraocular environment

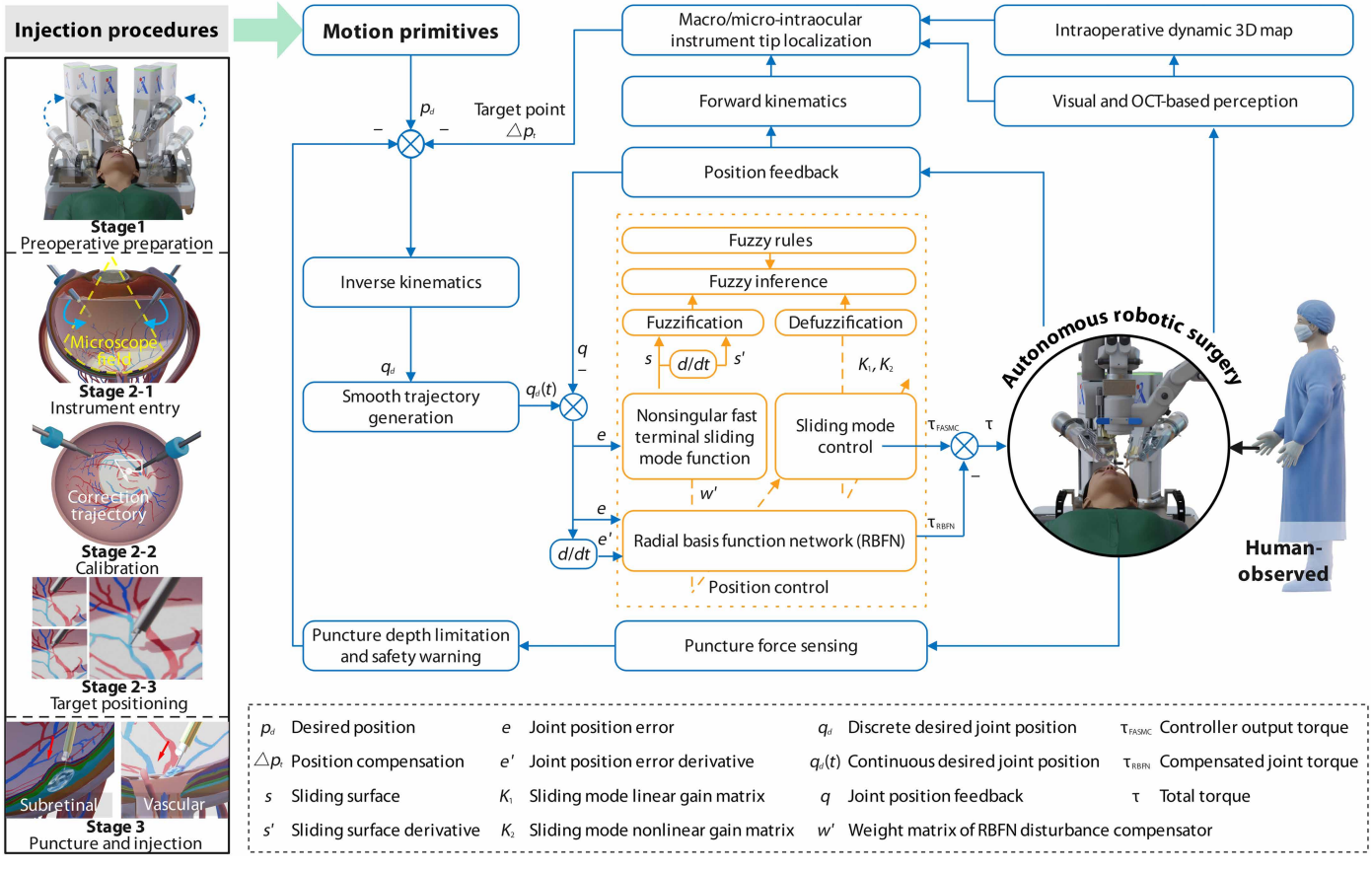


Fig. 4. Surgical procedure and unified control architecture for ARISE. The procedure consisted of three main stages. Stage 1 comprised ocular data acquisition and teleoperated insertion of dual-manipulator instruments. In stage 2, ARISE autonomously performed target localization, motion calibration, and positioning. Stage 3 executed autonomous retinal puncture for subretinal or vascular injection. A unified control architecture governed the entire procedure. Task-specific motion primitives were generated via task decomposition and were executed using inverse kinematics, trajectory smoothing, and a closed-loop controller based on fuzzy adaptive sliding mode control with a radial basis function network. Dynamic 3D mapping, macro/microcooperative localization, and real-time monitoring of instrument motion and puncture force ensured accuracy and safety throughout.

(Fig. 5A). Eight retinal patterns were derived from intraoperative fundus images of eight patients (Fig. 5B). For each pattern, 15 eyeball phantoms were fabricated, yielding a total of 120 phantoms ($n = 120$) for experimentation. Because of the absence of ocular circulation, the internal environment of the phantoms remained static. Although performed on phantoms, the autonomous workflow of ARISE strictly followed the same procedural steps as in real intraocular surgeries (Fig. 5C). After the surgeon teleoperatively inserted the instruments via the robotic dual manipulators under the microscope, the microscope's FOV was magnified to focus on the region typically observed during surgery. All subsequent steps were performed autonomously. In stage 2.1, the dual manipulators entered the visible field, followed by calibration initialization in stage 2.2. Stage 2.3 formed the core of the autonomous procedure, in which the injection needle approached and aligned with the target area. After the coarse alignment, fine adjustment was executed. Stage 3 involved the puncture, which was carried out after the alignment process. To evaluate positioning accuracy across varying instrument approach angles, the target injection area was divided into five regions centered around the optic disc (Fig. 5D). Regions 1 to 3 were within the vascular arch, with region 2 corresponding to the macular area, whereas regions 4 and 5 were outside the arch. In the absence of intraretinal vasculature, puncture depth in stage 3 was limited to retinal surface contact. Positioning errors along the x and y axes were measured from microscope images, and z-axis accuracy was determined by puncture force at the needle tip. An additional evaluation distance ensured puncture force generation, and an error was defined as the deviation between expected and actual puncture distances at the retinal contact point.

Experimental results demonstrated high positioning accuracy (Fig. 5D). Positioning accuracy was a key metric for precise targeting and for minimizing the risk of retinal injury, directly influencing surgical safety and effectiveness. The average positioning errors were 15.39 μm along the x axis, 15.59 μm along the y axis, and 3.92 μm along the z axis. The corresponding maximum errors were 45.75, 43.39, and 9.71 μm, respectively. In the simulated ocular environment, ARISE required on average 154.81 s to complete the full autonomous workflow, from instrument insertion to task completion. The maximum recorded duration was 209.1 s, indicating overall procedural efficiency within the constrained intraocular setting (Fig. 5E). Among all stages, calibration took the longest, whereas execution in region 2 was the fastest, because the calibrated position was typically located in this region, reducing travel distance.

Autonomous injection in ex vivo porcine eyeballs

To demonstrate the effectiveness of targeted retinal injection with ARISE, experiments were conducted on ex vivo porcine eyeballs with a fixed injection volume of 50 μl (Fig. 6A). The experiments included both subretinal and retinal vein injections. For the vascular injections, two target sites were selected, the CRV and BRV (*57*), because they represent distinct vascular regions. The CRV at the optic disc was larger and densely crowded, whereas the BRVs in the periphery were thinner and more sparsely distributed, and they were therefore analyzed separately. Ex vivo porcine eyeballs mounted on mannequin heads exhibited local dynamic responses to instrument contact, whereas the intraocular environment remained semistatic because of the absence of systemic physiological activity.

Experimental results demonstrated a 100% success rate for autonomous injections in ex vivo porcine eyeballs, including subretinal ($n = 20$), CRV ($n = 20$), and BRV ($n = 20$) injections. For each injection type ($n = 20$), the 95% confidence interval was 83.2%, 100.0%. As shown in Fig. 6B, a retinal bulge was observed after subretinal injection. Retinal arteries and veins originated from the optic disc and extended toward periphery. Consequently, during injection in the CRV region, blood flowed radially. During injection in the BRV, blood primarily flowed in the direction aligned with the needle orientation. In ex vivo porcine eyeballs, the average duration for ARISE to autonomously complete the surgery, from instrument insertion to completion, was 188.83 s. The maximum duration was 254.86 s, slightly longer than for the eyeball phantom (Fig. 6C). This increase was attributed to an additional 10 s allocated to injection and the need to expand the matching area. The expansion was necessitated by the more dispersed retinal features in the porcine eyeball, which in turn increased computational time.

ARISE demonstrated high positioning accuracy during autonomous injection in ex vivo porcine eyeballs (Fig. 6D). The average positioning errors were 16.29 μm (x axis), 14.25 μm (y axis), and 12.52 μm (z axis) for subretinal injections; 16.72 μm (x axis), 11.39 μm (y axis), and 3.27 μm (z axis) for CRV injections; and 17.89 μm (x axis), 13.78 μm (y axis), and 7.28 μm (z axis) for BRV injections. The maximum errors were 46.79 μm (x axis, BRV), 47.61 μm (y axis, BRV), and 26.99 μm (z axis, subretinal). During subretinal injection, the needle tip was required to reach the subretinal space to deliver the agent. This space was a potential gap located between the retinal pigment epithelium and the photoreceptor layer. The retinal pigment epithelium was typically located 45.8 ± 8.5 μm from the outer limiting membrane (OLM). Because of retinal adhesion in ex vivo porcine eyes, separating the retinal pigment epithelium from the photoreceptor layer was challenging. Therefore, the puncture depth was measured relative to the OLM, with a required downward penetration of 35.8 μm. During evaluation, the distance between the needle tip indentation and the OLM was assessed via iOCT imaging. For vascular injection, the optimal puncture site was at the center of the vessel, as determined by the cannula tip position. The puncture depth was assessed on the basis of the time interval between puncture and injection. The lower resolution of the iOCT B-scan compared with the A-scan contributed to the higher error observed in subretinal injection.

Autonomous injection in in vivo animal eyeballs

To further validate ARISE's surgical outcomes under fully dynamic ocular conditions, in vivo experiments were conducted on animal eyeballs, with a fixed injection volume of 50 μl (Fig. 7A). Bama miniature pigs and New Zealand White rabbits were used, with porcine eyeballs undergoing both subretinal and vascular injections and rabbit eyeballs limited to subretinal injections. Subretinal results from both species were analyzed jointly, because grouping was determined by injection targets rather than by animal species. In contrast with phantoms and ex vivo eyeballs, the intraocular environment was dynamic because of physiological activity. All animals were maintained under general anesthesia, ensuring that ocular motion remained within a controlled small amplitude.

The experimental results demonstrated that ARISE achieved a 100% success rate for autonomous injections in in vivo animal eyeballs despite interference from physiological motion, including subretinal ($n = 16$; 12 in porcine eyeballs and 4 in rabbit eyeballs), CRV ($n = 16$), and BRV ($n = 16$) injections. For each injection type ($n = 16$), the corresponding 95% confidence interval was 79.4%,

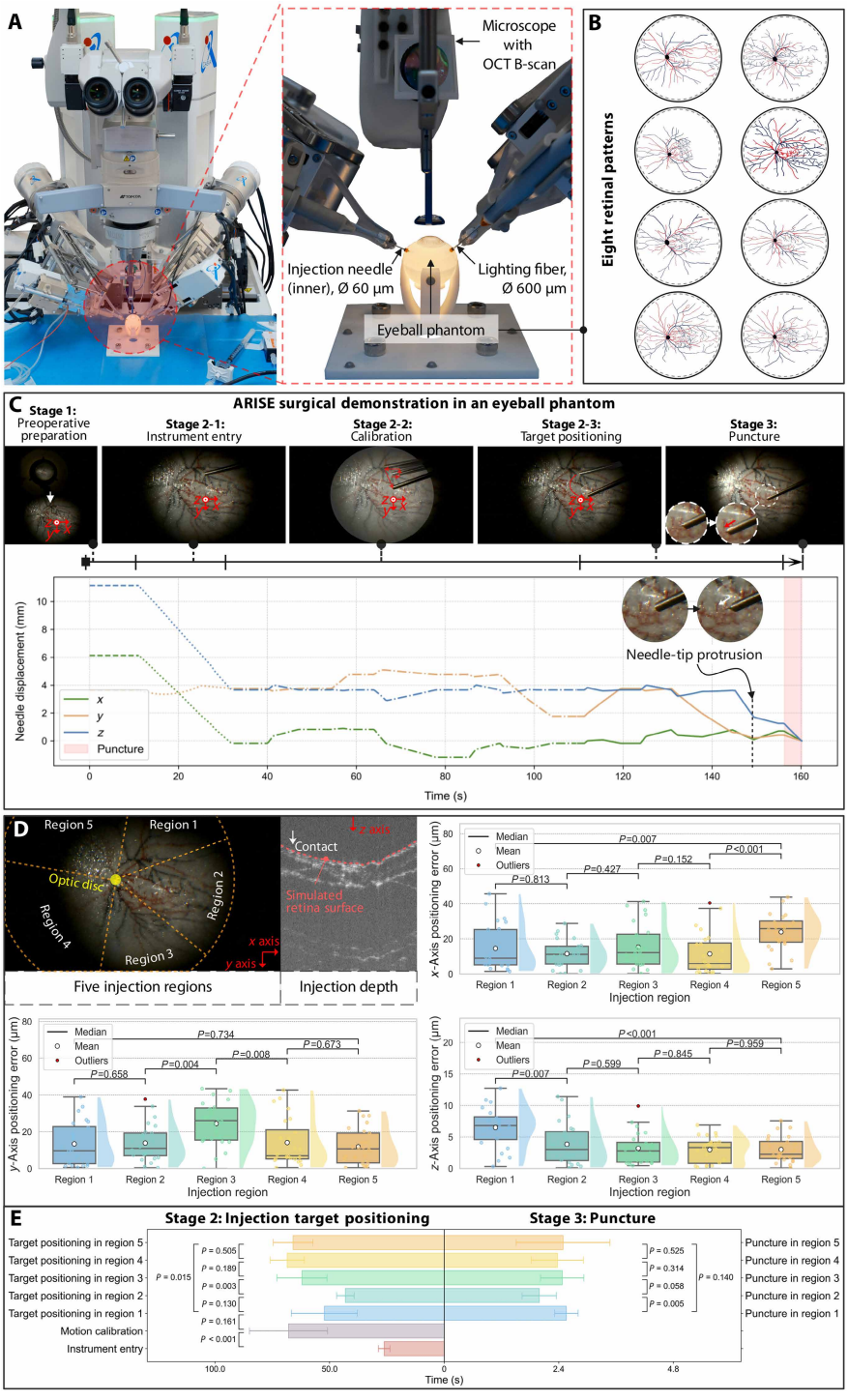


Fig. 5. Results of autonomous injection localization with eyeball phantoms. All experiments were performed under static intraocular conditions because of the absence of physiological motion and ocular circulation. (**A**) Experimental setup. (**B**) Eight retinal patterns, including optic disc and vessels. (**C**) Typical intraocular workflow performed by ARISE. (**D**) Regional division for injection targeting, puncture depth definition, and positioning error distribution. (**E**) Time required for each stage of the autonomous workflow.

100.0%. As shown in Fig. 7B, subretinal injections produced retinal bulges, CRV injections caused outward blood dispersion followed by vessel refilling, and BRV injections displaced blood that then flowed along the vein toward the optic disc. Under in vivo conditions, ARISE completed the full procedure autonomously. The average duration from instrument insertion to the final operation was 194.38 s, with a maximum recorded time of 271.68 s (Fig. 7C). To ensure puncture accuracy under in vivo conditions, ARISE maintained the same control strategy while requiring additional time for navigation, target localization, and error correction, which led to an extended puncture duration.

ARISE demonstrated high positioning accuracy during autonomous injection under in vivo conditions (Fig. 7D). The average positioning errors were 16.58 μm (*x* axis), 12.65 μm (*y* axis), and 9.30 μm (*z* axis) for subretinal injections; 17.44 μm (*x* axis), 10.26 μm (*y* axis), and 4.68 μm (*z* axis) for CRV injections; and 19.56 μm (*x* axis), 10.33 μm (*y* axis), and 6.22 μm (*z* axis) for BRV injections. The maximum errors were 53.17 μm (*x* axis, BRV), 33.13 μm (*y* axis, BRV), and 23.61 μm (*z* axis, subretinal). The puncture criteria were consistent with those applied in the ex vivo porcine eyeballs. During subretinal injection, prominent bulges above the OLM were observed intraoperatively. These bulges involved the inner segments and part of the outer segments and were consistent with the morphological criteria required for successful injection. During vascular injections, blood vessels in the in vivo porcine eyeballs were subsequently refilled, and iOCT was used to visualize the blood-flushing effect. Transient intravascular cavities were observed during the injections. Throughout all subretinal and vascular injections, no intraoperative or immediate postoperative complications, such as hemorrhage, retinal tears, or retinal detachment, were recorded. Furthermore, 1-week postoperative follow-up revealed no signs of visual impairment, ocular opacity, or abnormal animal behavior, indicating the minimally invasive nature of the intervention.

Comparison among manual, teleoperated, and autonomous injections

The effectiveness of ARISE in targeted retinal injection was evaluated through vascular and subretinal injections on ex vivo porcine eyeballs (Fig. 8A), in comparison with manual and teleoperated surgeries. Autonomous, manual, and teleoperated surgical modalities were each tested using an equal number of samples to allow a valid statistical comparison. Only vascular injection positioning accuracy was included in the error analysis, because human tracking of the target relied on prominent anatomical landmarks such as vascular junctions. The manually controlled injection instrument included iOCT, microforce sensing at the needle tip, and optical tracking. The same robotic platform was used for both teleoperated and autonomous procedures, ensuring consistency in the dual-manipulator instruments. During

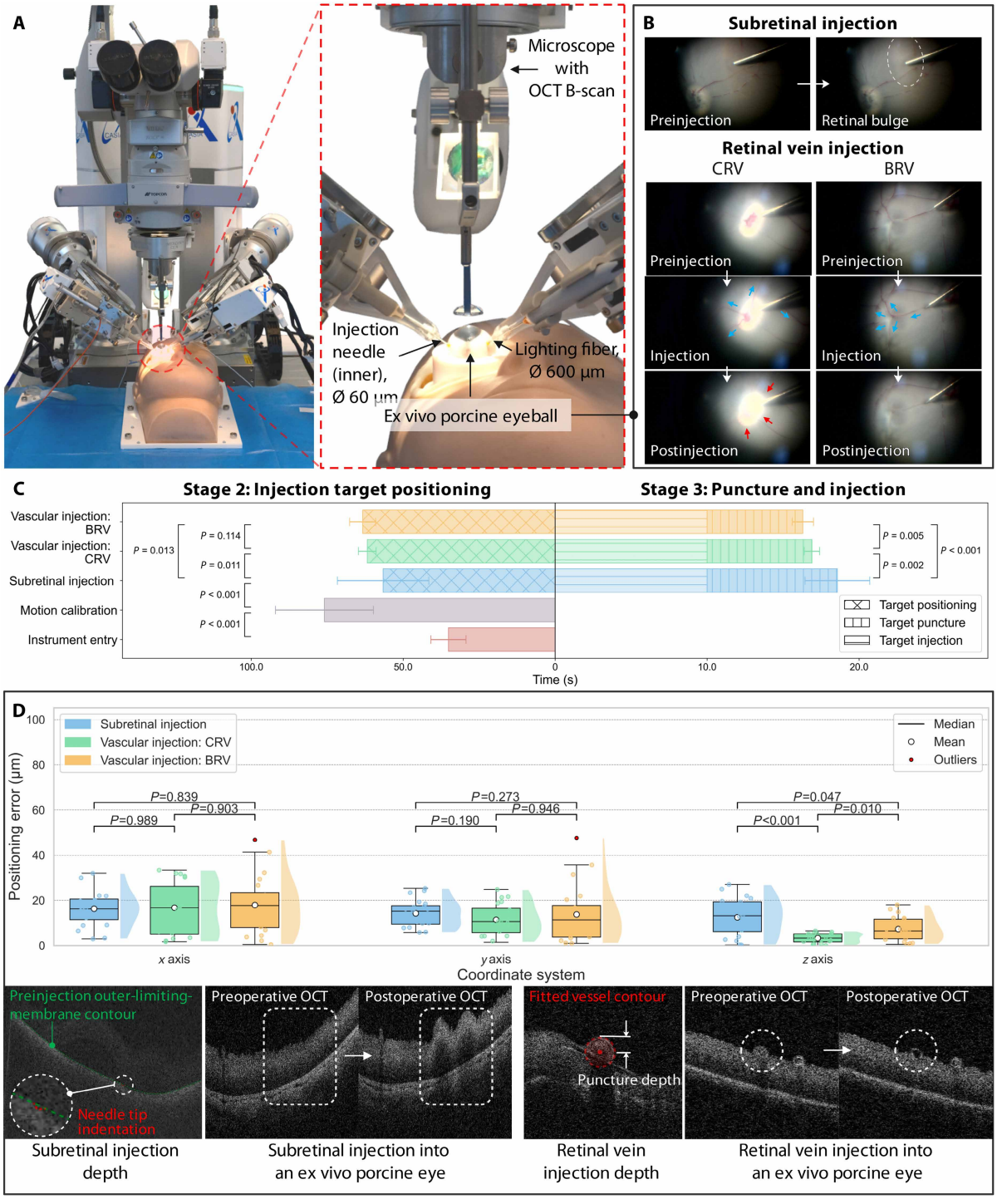


Fig. 6. Results of autonomous injection in ex vivo porcine eyeballs. All experiments were conducted under semistatic intraocular conditions, characterized by local tissue deformation, vitreous disturbance, and the absence of physiological motion such as respiration or blood flow. (**A**) Experimental setup. (**B**) Results of subretinal and retinal vein injections. (**C**) Time required for each stage of the autonomous workflow. (**D**) Positioning error, puncture depth, and the iOCT comparison before and after injection.

teleoperation, the system provided detailed navigation, including intraocular tissue layers, target-relative positioning, and puncture depth.

The motion trajectory of the injection needle tip during target positioning was analyzed on the basis of optical tracking data (Fig. 8B). Because of the 0.12-mm root mean square error in optical tracking, minor fluctuations were observed even during stable robotic surgery. In autonomous surgery, fluctuations were primarily concentrated in the latter half of the task. During this stage, fine-tuning involved adjustments on the scale of several tens of micrometers. These small adjustments were amplified by the tracking system,

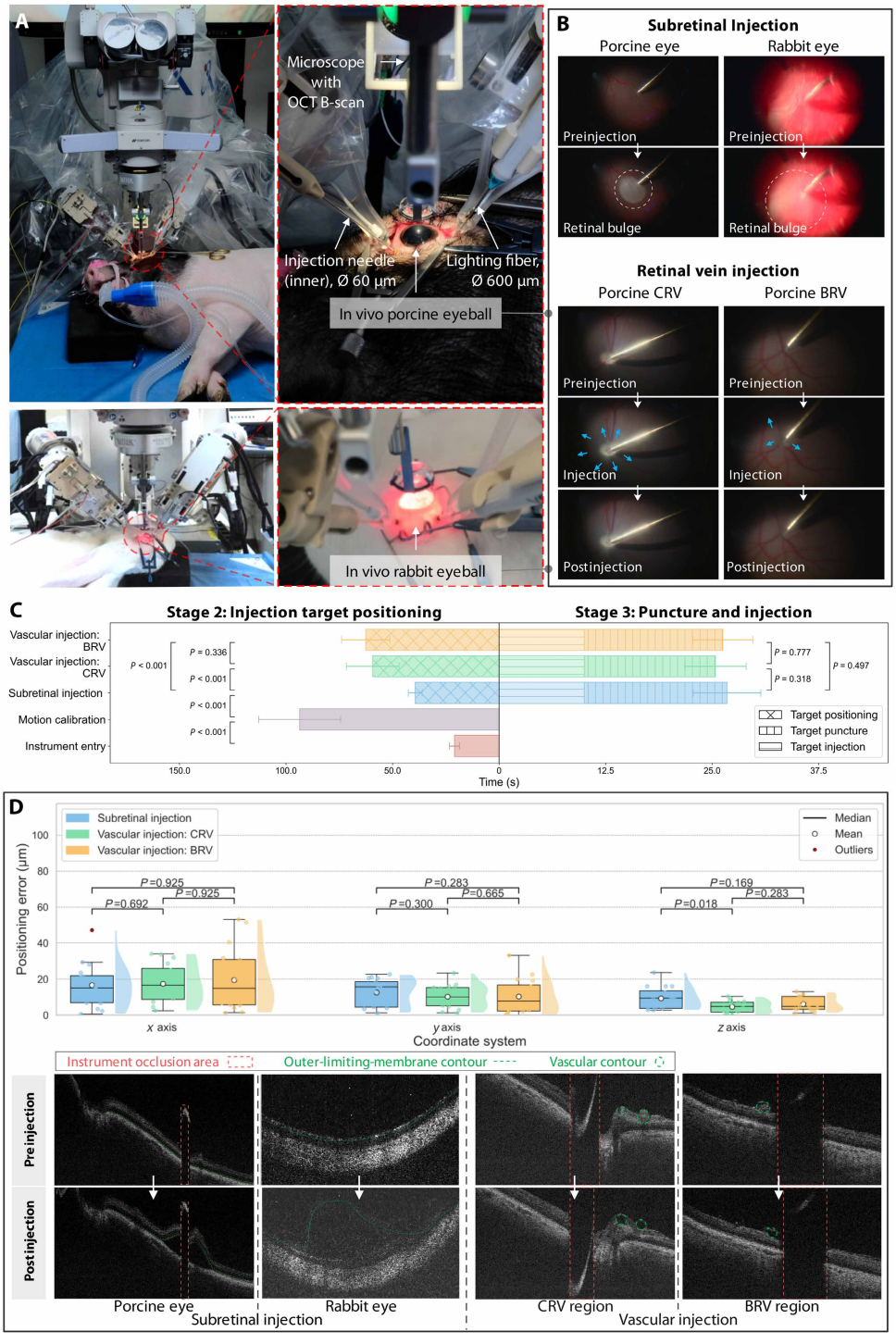


Fig. 7. Results of autonomous injection in in vivo animal eyeballs. All experiments were conducted under fully dynamic intraocular conditions, including physiological motion such as respiration and blood flow. (**A**) Experimental setup. (**B**) Results of subretinal and retinal vein injections. (**C**) Time required for each stage of the autonomous workflow. (**D**) Positioning error, puncture depth, and the iOCT comparison before and after injection.

resulting in visible motion deviations. In teleoperated robotic surgery, fluctuations persisted throughout the entire process, because the instrument tip did not maintain a constant velocity like in autonomous control. Because pitch, yaw, and feed of the injection needle were manually controlled during teleoperation, fluctuations were particularly pronounced along the *z* axis. Manual surgery exhibited substantially greater jitter because of natural hand tremor, which was further amplified by the tracking error. In contrast, autonomous surgery showed improved motion stability, consistent speed, reduced fluctuations, and greater safety compared with both teleoperated and manual surgeries.

Motion trajectories of the injection needle tip during equal-distance puncture were analyzed on the basis of iOCT A-scan signals (Fig. 8C). Manual puncture exhibited the shortest duration, although it lacked positional stability, which led to repeated attempts within the soft tissue. During teleoperated surgery, occasional pauses occurred during puncture, producing a stepped trajectory. In contrast, the puncture process in autonomous surgery was stable and maintained a consistent speed.

Autonomous and teleoperated robotic surgeries both achieved a 100% injection success rate, indicating precise and reliable performance. A successful injection involved a single-attempt puncture of the target site and precise delivery of the agent into the subretinal space or vessel lumen. Successful injection was confirmed by distinctly observable target-specific responses, such as bleb formation or blood displacement. It was also characterized by the absence of complications, including hemorrhage, retinal tears, or detachment. Because of inherent physiological tremor, manual vascular injection rarely reached 100% success, particularly in small-diameter vessels within the BRV (Fig. 8D).

The additional time required for autonomous surgery, which averaged 56.68 s more than manual surgery (Fig. 8E), remained well within the acceptable margin relative to the total surgery duration of ~30 min. The average duration of the stage from target positioning to injection was 79.16 s for autonomous surgery, 53.78 s for teleoperated surgery, and 22.48 s for manual surgery. By comparison, preoperative vitreous removal remained the most time-consuming step in clinical surgery, typically requiring 184 ± 56.9 s (*58*).

Autonomous robotic surgery exhibited the highest positioning accuracy (Fig. 8F), with an average error of 11.71 μm compared with 25.80 μm for teleoperated robotic surgery and 58.17 μm for manual surgery. The average positioning errors along the *x* axis were 17.31 μm for autonomous surgery, 33.02 μm for teleoperated robotic surgery,

and 55.40 μm for manual surgery. The average positioning errors along the *y* axis were 12.59 μm for autonomous surgery, 21.71 μm for teleoperated robotic surgery, and 64.93 μm for manual surgery. The average positioning errors along the *z* axis were 5.23 μm for autonomous surgery, 22.67 μm for teleoperated robotic surgery, and 54.19 μm for manual surgery. Manual surgery exhibited the greatest positioning error, which was primarily caused by natural tremors in the human hand. Because of the slight delay in human response, the *z*-axis positioning error in teleoperated robotic surgery was greater than in autonomous surgery despite depth information provided by the navigation.

DISCUSSION

This study introduced ARISE, a fully autonomous robotic platform developed for targeted retinal injection. A dual-manipulator microsurgical robot with high precision and dexterity was developed. ARISE's comprehensive navigation module integrated multiview spatial fusion, which constructed an intraoperative dynamic global 3D map for comprehensive intraocular perception. Criterion-weighted fusion of multisensor data enabled precise macro/micropositioning of the robotic instrument tip across the intraocular region. Multiconstraint objective optimization determined the trajectory of the robotic end effector by incorporating multiple intraoperative constraints, whereas safety was ensured through human-observed force-position-imaging-hybrid control. Clinical feasibility was validated using three experimental models: eyeball phantoms, ex vivo porcine eyeballs, and in vivo rabbit and porcine eyeballs. Validation was performed by comparing outcomes with both manual and teleoperated surgeries performed by surgeons. ARISE showed improved safety and positioning accuracy compared with manual and teleoperated approaches, particularly in challenging vascular injections. Robotic autonomy ensured consistent procedural accuracy and stability, minimizing human errors. Automated workflows reduced surgical workload, enabling greater focus on diagnostic and planning tasks. In the future, ARISE will be further developed for broader application across intraocular surgeries, enhancing both treatment diversity and accessibility.

Despite the demonstrated capabilities of ARISE, key challenges remain for its clinical translation, including system generalizability, tissue motion and deformation, pathological evaluation, and regulatory readiness. First, the generalizability of autonomous intraocular systems remains a core challenge. Existing platforms often rely on limited visual fields, relative localization, and manual control, restricting their consistency and scalability. ARISE enabled global intraocular navigation without requiring constant tool visibility, representing a notable advance. However, such autonomy must remain robust across diverse anatomies. In practice, sensor fusion may degrade under low signal-to-noise ratios. Although ARISE incorporated learning-based perception and outperformed previous systems in navigation capabilities, its performance under pathological conditions has not yet been fully validated. Future work should enhance the system's adaptability and uncertainty tolerance to ensure reliable operation in varied clinical scenarios. Second, intraoperative tissue motion and deformation remain major challenges for autonomous precision. On the one hand, despite controlled ventilation anesthesia, micrometer-scale ocular micromotions from respiration and blood flow pulse persist in vivo, affecting needle stability during prolonged procedures. Recent studies, such as that by Posselli *et al.* (*10*), have begun exploring teleoperated surgery under natural eye motion, highlighting potential pathways toward systems capable of accommodating physiological motion. On the other hand, interactions between instruments and delicate intraocular tissues introduce biomechanical uncertainties, including prepuncture deformation, elastic rebound, and lateral displacement. Although the application of ultrafine instruments and integrated force sensing helps mitigate these effects, broader tasks involving larger tools such as forceps or retractors may induce greater tissue deformation. Extending autonomy to such cases will require high-frequency, high-accuracy modeling and compensation to preserve micrometer-level precision and safety. Third, in terms of clinical evaluation, the present study did not involve disease modeling and thus lacks assessment of pathological factors that may influence intraoperative injection behavior, particularly drug reflux and postoperative tissue response. Although it plays a critical role in assessing delivery precision, minimal vitreous reflux is clinically acceptable. Currently, no standardized methods exist for quantifying reflux volume in either clinical or experimental settings. Although 1-week postinjection safety data in healthy animals revealed no ocular complications, these findings do not reflect pathological tissue complexity or long-term outcomes. Future studies will incorporate disease-specific models and extend postoperative monitoring to systematically evaluate drug retention, tissue response, and therapeutic efficacy under clinically relevant conditions. Separating results by species and vascular subtype may also yield additional insights when combined with disease modeling, and future studies will refine grouping strategies accordingly. Fourth, the autonomous nature of ARISE poses critical challenges for clinical translation, stemming from the shift in decision-making from surgeon supervision to machine autonomy. Unlike existing ophthalmic robots that rely on teleoperation, ARISE executes fully autonomous intraocular navigation and control, necessitating rigorous validation of its accuracy and safety under clinical uncertainties. To this end, a phased strategy is adopted involving modular validation within clinically approved teleoperated systems, enabling systematic refinement of autonomous functions. Clinical deployment will further require industrial collaboration for productization, adherence to internationally recognized manufacturing standards such as ISO 13485, and regulatory approval through national authorities. As autonomous systems extend beyond current regulatory paradigms, substantial efforts must also be devoted to establishing standards for interpretability, auditability, and ethical accountability to ensure clinical and regulatory acceptance.

MATERIALS AND METHODS

Dual-manipulator microsurgical robot

The dual-manipulator microsurgical robot was independently developed by the research team. It featured a coaxial orthogonal mirrored configuration and included lightweight manipulators fabricated from composite materials to enhance flexibility. Damping mechanisms were incorporated to suppress microvibrations originating from the motors and transmission elements. The control module used EtherCAT communication for high-speed and reliable data transmission among force sensors, servo controllers, and input/output modules. Kinematic errors were compensated by modeling backlash and transmission deviations in the harmonic reducer to improve positioning accuracy. A nonlinear compensation model was applied to correct dynamic errors resulting from joint friction and stiffness, thereby reducing delay-loop effects and improving control stability. The robot featured a sealed isolation structure and a sterile

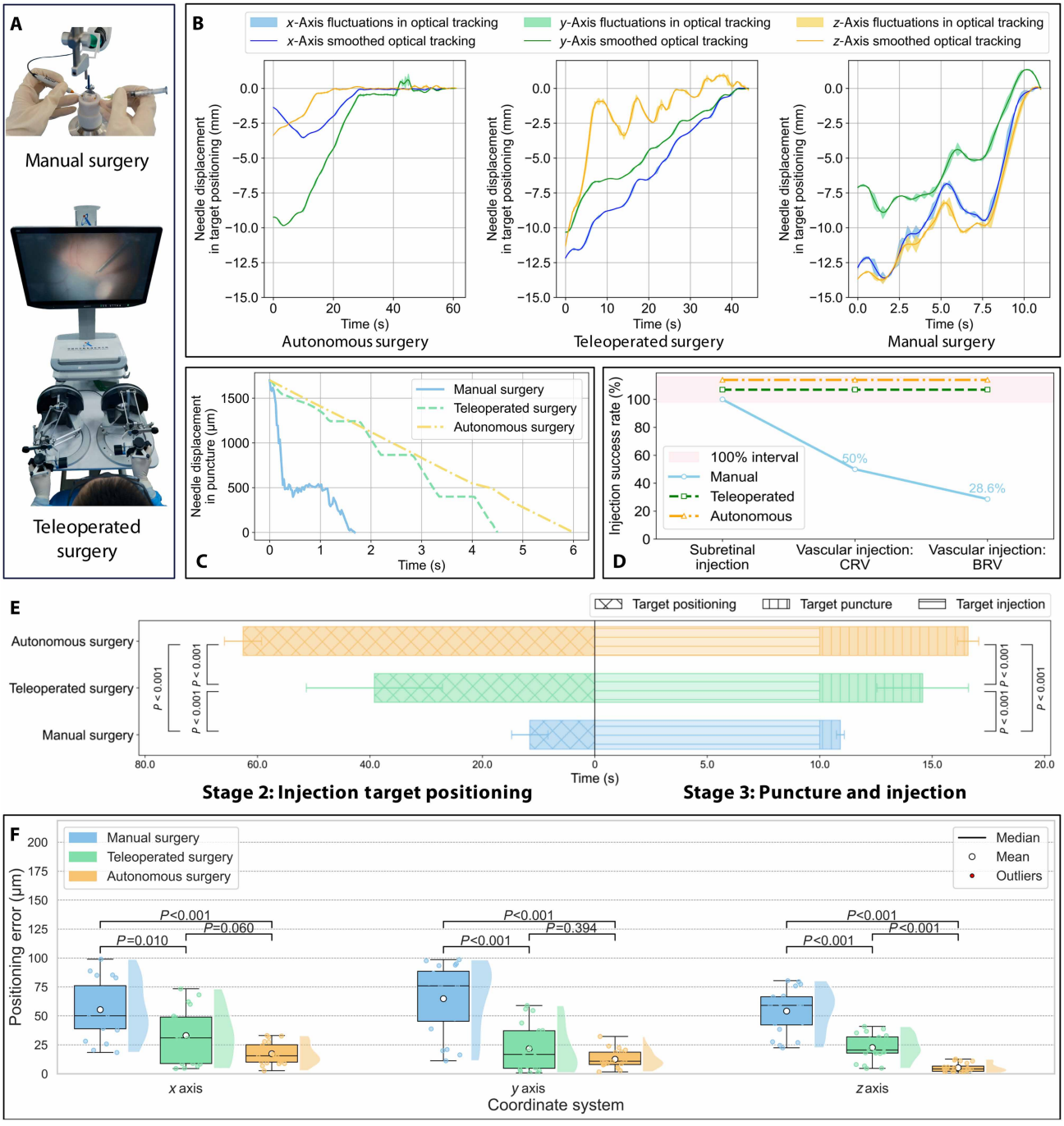


Fig. 8. Comparison of autonomous, manual, and teleoperated surgeries in ex vivo porcine eyeballs. (**A**) Experimental setup. (**B**) Motion trajectories of the injection needle tip during target positioning. (**C**) Comparison of motion trajectories over an equal puncture distance. (**D**) Injection success rates. (**E**) Comparison of time consumption across the same surgical stage. (**F**) Comparison of positioning error across autonomous, teleoperated, and manual surgeries.

protective cover, designed to comply with surgical sterilization and asepsis standards. The right end effector was an independently developed design that integrated a straight needle, a drug injection module, a linear motor, and a servo motor. The left end effector was fitted with a 23-gauge (~0.573 mm in diameter) medical-grade optical fiber (Alcon, USA), designed for intraocular surgical tasks.

Multimodal optical data setup

To support accurate intraoperative navigation and real-time safety monitoring, a multimodal optical data setup was constructed. It combined iOCT signals, preoperative OCT volumes and fundus photographs, microscope images, and extraocular tracking data. Among these components, a high-frequency, high-sensitivity iOCT

device was independently developed to provide perspective imaging during surgery. The iOCT device integrated a B-scan optical path shared with the surgical microscope and an A-scan probe embedded at the tip of the injection needle. This configuration enabled high-resolution 3D visualization of intraocular tissue. It also allowed localized depth-layer measurements from the needle tip to the retina. The diameter of the A-scan probe was 168 μm. This imaging unit provided an axial FOV of 3.62 mm, an axial resolution of 5.3 μm in tissue, an A-scan rate of 100.16 kHz, and a B-scan lateral imaging range of 16 mm by 16 mm. Preoperative imaging was conducted using the OCT device (3D OCT-1 Maestro, TOPCON, Japan), which captured OCT volumes and fundus photographs for navigation initialization. It generated a 3D retinal map with a 12 mm–by–9 mm lateral FOV, a 9-mm depth, a lateral resolution of 20 μm, and an axial resolution of 6 μm. During surgery, the surgical microscope (OMS-800, TOPCON, Japan), equipped with a 120D wide-angle lens and a 130° FOV, captured images at a 2560 pixel–by–1440 pixel resolution. To track spatial positioning during surgery, the optical tracker (Polaris Vega XT, Northern Digital Inc., Canada) monitored both the microscope's FOV center and end-effector positions of the dual-manipulator robotic instruments. Reflective marker tools were mounted on extraocular components. The setup avoided direct near-infrared irradiation of the eyeball at a close range, preventing errors caused by marker displacement or light exposure. According to the manufacturer's specifications, the optical tracker achieved a volumetric accuracy of 0.12 mm (root mean square), with a 95% confidence interval root mean square of 0.20 mm. It operated at 60 Hz with a latency below 16 ms.

Coarse intraocular reconstruction based on pupil-trocar geometry

Coarse intraocular reconstruction based on pupil-trocar geometry constructed a personalized intraocular geometric model for ARISE navigation. Inputs included instrument pose measurements from the optical tracker (Polaris Vega XT, Northern Digital Inc., Canada) and mean anatomical parameters of the eyeball. A global coordinate frame was defined by aligning the z axis with the microscope optical axis, with the x and y axes in the orthogonal plane. The pupil center was initialized from the microscope FOV center tracked optically, and an eye model was instantiated from human, pig, and rabbit anatomical datasets. During trocar insertion, dual-manipulator robot poses were recorded to locate trocar positions in the global frame, coinciding with the robot RCM. According to medical standards, trocar sites on the pars plana were symmetrically placed at fixed distances from the corneal limbus. These constraints ensured that dual-arm entry points lay on the same circumference, enabling estimation of eyeball orientation. Roll around the ocular axis was disregarded because of the spherical approximation of the eyeball. The output was an initialized intraocular model encoding spatial relationships among the microscope, robot, and eyeball, serving as the foundation for retinal structure modeling.

Registration of global microfeatures in fundus images

The registration of global microfeatures in fundus images registered horizontal retinal microfeatures to mitigate boundary ambiguity and occlusion, providing stable in-plane landmarks for ARISE navigation. It integrated intraoperative microscope and preoperative retinal color images. A multiscale, two-stage matching strategy combined sparse and dense correspondences. Microfeatures were extracted from segmented images. Coarse matches were established with SuperPoint-SuperGlue and LoFTR, and DBSCAN clusters refined covisible regions. GlueStick improved precision by leveraging local structures. Match shifts were corrected by removing keypoints on instruments and aligning valid trajectories across frames, ensuring that correspondences remained on the retinal surface. RANSAC (random sample consensus)–based filtering and homography estimation fused pre- and intraoperative images for consistent alignment. The output achieved 90.47% accuracy in refined horizontal feature detection. It recovered 99.70% of features in instrument-obscured regions and 97.47% in invisible areas through preoperative and intraoperative matching. These features were integrated into the 3D retinal map to recover missing details and support augmented-reality visualization.

Lateral field feature mapping–based 3D fundus image registration algorithm

The lateral field feature mapping–based 3D fundus image registration algorithm fused layered information from preoperative OCT, preoperative retinal images, and intraoperative iOCT B-scans to refine retinal modeling. It also incorporated horizontal correspondences from the registration of global microfeatures in fundus images. For phantom experiments, preoperative OCT and fundus images were taken from patient scans used in model construction. For ex vivo porcine and in vivo animal eyeballs, OCT and fundus data were acquired with the OCT device (3D OCT-1 Maestro, TOPCON, Japan). Retinal contours were segmented from OCT using ResUNet (*59*). Correspondences were established between B-scan slices and horizontal features from microscope images. These guided integration of OCT-derived axial information with lateral alignment. A 3D deformation matrix compensated for the axial displacement of the intraoperative fundus induced by eyeball rotation. The output was an axially corrected 3D retinal reconstruction supporting accurate intraoperative navigation.

Extrinsic spatial localization assisted by the RCM with a 100-μm accuracy

The extrinsic spatial localization algorithm, assisted by the RCM, estimated instrument tips outside the microscope's FOV by fusing optical tracking and robot forward kinematics. Reflective markers on the instrument were measured by the optical tracker (Polaris Vega XT, Northern Digital Inc., Canada), yielding global tip localization but with reduced accuracy due to the marker error (0.12 mm), tip-marker offsets, and occlusion. Forward kinematics exploited the RCM mechanism of the ophthalmic robot, where the trocar acted as the fixed center of yaw, pitch, and feed joints. This yielded a high relative accuracy (7 arc sec for yaw/pitch and 0.005 mm for feed), allowing micrometer-precision tip coordinates, but only in local coordinates. The extrinsic spatial localization algorithm aligned robot and microscope frames through the RCM. A transformation matrix mapped high-precision kinematic coordinates into the global frame. Dynamic weighting between the two sources compensated for angular deviations, trocar-induced scleral deformation, and cumulative errors. The fusion integrated coarse global tracking with precise kinematics, yielding globally referenced tip localization with 100-μm global accuracy and micrometer-level relative precision.

End-effector localization algorithm with a sub–10-μm accuracy

The end-effector localization algorithm operated when the instrument entered the posterior segment and became visible under the

microscope. SurgiNet (*45*) segmented the injection needle, illumination fiber, and shadows with a Dice score of 92.35%. Central axes were extracted from contours to model the end effector and enabled lateral localization. Axial localization was derived from the geometry of the light source, needle tip, and shadow projection. The true source-tip distance was obtained from robot kinematics, whereas lateral displacement was measured from image localization. Similar-triangle analysis yielded the tip-shadow depth. By referencing the global retinal map, the axial offset between the projected shadow tip and the injection target was derived, providing the global z coordinate of the target. This projection-based method enabled full 3D localization of the end effector with a sub–10-μm accuracy.

iOCT-based relative localization with micrometer-level accuracy

The iOCT-based relative localization algorithm refined localization when the instrument approached the retinal surface and entered the iOCT range. Inputs included iOCT A-scan depth data and instrument-retina relationships derived from the end-effector localization algorithm and the global map. The axial distance between the needle tip and the retina was measured from A-scans with a 5.3-μm accuracy. Relative orientation was determined by decomposing the instrument-retina angular relationship, reconstructing the end-effector pose relative to the injection target. To integrate iOCT-based localization into the global frame, a 3D rotational transformation converted tip coordinates from the iOCT to the global frame, whereas the end-effector localization algorithm results maintained cross-modality consistency. The output achieved micrometer-level accuracy in position and orientation, enabling precise registration of the needle tip with the retinal target and surrounding tissue.

Control architecture

The control architecture was designed to execute complex surgical tasks with high precision and adaptability. This was achieved through hierarchical coordination of motion generation, disturbance rejection, and feedback mechanisms, as illustrated in Fig. 4. Task decomposition was performed via multiconstraint optimization. An improved rapidly exploring random tree algorithm generated motion primitives, including illumination adjustment, alignment, protrusion, puncture, and injection. Multiview spatial fusion was applied to multimodal intraoperative data to construct a dynamic 3D map, enabling flexible target tracking. Real-time localization of the end effector ensured precise positioning of the instrument tip. These target poses were sent to the low-level controller, which applied fifth-order polynomial interpolation and inverse kinematics to generate smooth joint-space trajectories. To maintain tracking accuracy under nonlinear disturbances and constraints, a fuzzy adaptive sliding mode controller with a radial basis function network enabled closed-loop control. A human-observed module monitored instrument motion using a domain knowledge–driven perception method (*60*). It detected puncture events through low-latency force sensing and provided real-time visualization and risk alerts to support safe, interruptible operation.

Eyeball phantom studies

Anatomically realistic eyeball phantoms were developed to validate intraocular surgery. The model size and shell followed Gullstrand Eye parameters as a general anatomical reference (fig. S1). Fundus patterns were generated from preoperative imaging data of eight patients, including macular hole and epiretinal membrane, to represent common retinal abnormalities. A 0° disposable contact lens simulated the cornea. The globe and fundus were molded with silicone of Shore A hardness 85 and 30, respectively. The retinal pattern layer was embedded between two silicone shells with a transparent PVC adhesive film. Inner retinal contours were reconstructed from OCT B-scans by surface interpolation. Fundus patterns were obtained by planar projection of fused scanning laser ophthalmoscopy and fundus photographs. The resulting pattern was affixed to the inner layer using a globe-mapping method. Figure S1B shows retinal surface fabrication. Figure S1C shows fundus assembly. For experiments, specimens were randomly assigned to groups using a computer-generated sequence, with allocation conducted independently of the investigators. Phantoms were mounted without rigid fixation to the orbital base. This allowed passive rotation around their geometric centers under external forces such as scleral trocar traction.

Ex vivo animal studies

Ex vivo porcine eye experiments were conducted to evaluate surgical performance under realistic anatomical and biomechanical conditions. Pigs are widely adopted as biological models in ophthalmic research (*61*) because their eyes resemble human eyes in globe size, scleral thickness, and retinal morphology. Freshly enucleated porcine eyeballs were procured at 6:00 a.m. from a certified animal supplier on each experimental day and kept refrigerated for surgery between 9:00 a.m. and 5:00 p.m. To minimize bias, specimens were randomly assigned to experimental conditions using a computer-generated sequence, and the allocation process was conducted independently of the investigators performing the experiments. The internal ocular structures, including the vitreous body, remained intact without retinal detachment or lens opacity, and the globe was not incised during surgery. Each eyeball was mounted on a standard ophthalmic surgical training mannequin head, which allowed free rotation with constrained pitch and yaw movements to simulate physiological eye motion, as shown in fig. S2A. The ex vivo eyeballs exhibited local dynamic responses, including soft tissue deformation, vessel displacement, and vitreous disturbance from instrument contact, and thus remained semistatic.

In vivo animal studies

In vivo validation was conducted on Bama miniature pigs (3 to 5 months, 15 to 25 kg) and New Zealand White rabbits (1.5 to 2.0 kg), as shown in fig. S2 (B and C), to assess the feasibility and safety of retinal injection under physiological conditions. Rabbits were widely used in biomedical research because their retinas contained a high density of photoreceptors, but their relatively thin and largely avascular structure was a limitation (*62*). Porcine eyeball experiments were the primary focus, involving both subretinal and vascular injections, whereas rabbits underwent only subretinal injections without randomization. Pigs were randomly assigned to experimental conditions using a computer-generated sequence, and the allocation process was conducted independently of the investigators conducting the experiments. All animals were housed under the Chinese national standard "Laboratory Animal—Requirements of Environment and Housing Facilities" (GB 14925-2010). Experimental protocols and procedures were approved by the Institutional Animal Care and Use Committee of Peking Union Medical College Hospital (XHDW-2023-092 for rabbits; XHDW-2023-120 for pigs). Unlike eyeball phantoms and ex vivo models, the intraocular environment in vivo was dynamic, influenced by respiration, cardiac motion, and perfusion. To

reproduce ocular motion in vitreoretinal surgery, animals were fasted for 12 to 24 hours and anesthetized under general anesthesia. For pigs, anesthesia was induced with intramuscular Su-Mian-Xin II (2 mg/kg; a xylazine hydrochloride–based agent), followed by orotracheal intubation and maintenance with a closed-circuit inhalation setup. Ventilation parameters were a tidal volume of 10 ml/kg, a respiratory rate of 20 breaths/min, an oxygen flow of 1 liter/min, an inspiration-to-expiration ratio of 2:1, and 2% isoflurane. For rabbits, anesthesia was induced by intravenous sodium pentobarbital (30 mg/kg) and intramuscular Su-Mian-Xin II (4 mg/kg; a xylazine hydrochloride–based agent). Both species were maintained under general anesthesia, which stabilized ocular motion. A previous study reported such motion to remain within an amplitude of <0.01 mm (*63*), approximating micromovements of the human eye during surgery. After induction, mydriasis was achieved with topical tropicamide. During surgery, the body temperature was maintained at 24° to 26°C with a heating pad, and artificial tears were applied to keep the cornea hydrated.

Statistical analysis

Statistical analyses were primarily performed with the SciPy library (version 1.10.0) in Python. The Clopper-Pearson exact method was used to compute confidence intervals for injection success rates (binary outcomes), which yielded robust estimates in small-sample scenarios. Intergroup comparisons of surgical performance metrics were conducted using the Mann-Whitney *U* test. These metrics included procedure duration across distinct surgical stages and positioning errors under varying surgical modalities and control strategies. This nonparametric method was suitable for data that did not conform to a normal distribution. Statistical significance was defined as $P < 0.05$. To ensure objectivity, a double-blind strategy was implemented for data processing in all experiments. Specifically, the personnel responsible for performing the surgeries were distinct from those conducting data processing and statistical analysis, and all datasets were anonymized before analysis to eliminate potential bias arising from experimental grouping or outcome expectations.

In the figures, box-and-whisker plots show the interquartile ranges [IQRs, quartile 1 (Q1) to Q3] as boxes, with the medians shown by horizontal lines and whiskers extending to the most extreme values within 1.5× IQR. Means are marked by white circles with black outlines, and outliers are plotted as colored circles. Horizontal bars depict the mean times for each stage, and error bars in bar graphs indicate ±1 SD.

Supplementary Materials

The PDF file includes:
Methods
Figs. S1 to S13
Tables S1 to S3

Other Supplementary Material for this manuscript includes the following:
Movie S1
MDAR Reproducibility Checklist

REFERENCES AND NOTES

1. L. Gu, S. Poddar, Y. Lin, Z. Long, D. Zhang, Q. Zhang, L. Shu, X. Qiu, M. Kam, A. Javey, Z. Fan, A biomimetic eye with a hemispherical perovskite nanowire array retina. *Nature* **581**, 278–282 (2020).
2. T. Qiu, Q. An, J. Wang, J. Wang, C.-W. Qiu, S. Li, H. Lv, M. Cai, J. Wang, L. Cong, S. Qu, Vision-driven metasurfaces for perception enhancement. *Nat. Commun.* **15**, 1631 (2024).
3. J. Meng, T. Wang, H. Zhu, L. Ji, W. Bao, P. Zhou, L. Chen, Q.-Q. Sun, D. W. Zhang, Integrated in-sensor computing optoelectronic device for environment-adaptable artificial retina perception application. *Nano Lett.* **22**, 81–89 (2021).
4. T. F. Tan, A. J. Thirunavukarasu, L. Jin, J. Lim, S. Poh, Z. L. Teo, M. Ang, R. V. P. Chan, J. Ong, A. Turner, J. Karlström, T. Y. Wong, J. Stern, D. S.-W. Ting, Artificial intelligence and digital health in global eye health: Opportunities and challenges. *Lancet Glob. Health* **11**, e1432–e1443 (2023).
5. D. Holmes, Reconstructing the retina. *Nature* **561**, S2–S3 (2018).
6. J. E. Dowling, Restoring vision to the blind. *Science* **368**, 827–828 (2020).
7. L. Ma, B. Fei, Comprehensive review of surgical microscopes: Technology development and medical applications. *J. Biomed. Opt.* **26**, 010901 (2021).
8. J. B. Lin, R. Narayanan, E. Philippakis, Y. Yonekawa, R. S. Apte, Retinal detachment. *Nat. Rev. Dis. Primers* **10**, 18 (2024).
9. D. Goldenberg, J. Shahar, A. Loewenstein, M. Goldstein, Diameters of retinal blood vessels in a healthy cohort as measured by spectral domain optical coherence tomography. *Retina* **33**, 1888–1894 (2013).
10. N. R. Posselli, E. S. Hwang, Z. J. Olson, A. Nagiel, P. S. Bernstein, J. J. Abbott, Head-mounted surgical robots are an enabling technology for subretinal injections. *Sci. Robot.* **10**, eadp7700 (2025).
11. I. I. Iordachita, M. D. De Smet, G. Naus, M. Mitsuishi, C. N. Riviere, Robotic assistance for intraocular microsurgery: Challenges and perspectives. *Proc. IEEE* **110**, 893–908 (2022).
12. E. V. Poorten, C. N. Riviere, J. J. Abbott, C. Bergeles, M. A. Nasseri, J. U. Kang, R. Sznitman, K. Faridpooya, I. Iordachita, "36 - Robotic retinal surgery" in *Handbook of Robotic and Image-Guided Surgery*, M. H. Abedin-Nasab, Ed. (Elsevier, 2020), pp. 627–672.
13. R. Ladha, T. Meenink, J. Smit, M. D. De Smet, Advantages of robotic assistance over a manual approach in simulated subretinal injections and its relevance for gene therapy. *Gene Ther.* **30**, 264–270 (2021).
14. K. Willekens, A. Gijbels, L. Schoevaerdts, L. Esteveny, T. Janssens, B. Jonckx, J. H. M. Feyen, C. Meers, D. Reynaerts, E. V. Poorten, P. Stalmans, Robot-assisted retinal vein cannulation in an in vivo porcine retinal vein occlusion model. *Acta Ophthalmol.* **95**, 270–275 (2017).
15. A. Gijbels, J. Smits, L. Schoevaerdts, K. Willekens, E. B. V. Poorten, P. Stalmans, D. Reynaerts, In-human robot-assisted retinal vein cannulation, a world first. *Ann. Biomed. Eng.* **46**, 1676–1685 (2018).
16. J. Smits, M. Ourak, A. Gijbels, L. Esteveny, G. Borghesan, L. Schoevaerdts, K. Willekens, P. Stalmans, E. Lankenau, H. Schulz-Hildebrandt, G. Huttmann, D. Reynaerts, E. B. V. Poorten, "Development and experimental validation of a combined FBG force and OCT distance sensing needle for robot-assisted retinal vein cannulation," in *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2018), pp. 129–134.
17. T. L. Edwards, K. Xue, H. C. M. Meenink, M. J. Beelen, G. J. L. Naus, M. P. Simunovic, M. Latasiewicz, A. D. Farmery, M. D. De Smet, R. E. MacLaren, First-in-human study of the safety and viability of intraocular robotic surgery. *Nat. Biomed. Eng.* **2**, 649–656 (2018).
18. J. Cehajic-Kapetanovic, K. Xue, T. L. Edwards, T. C. Meenink, M. J. Beelen, G. J. Naus, M. D. De Smet, R. E. MacLaren, First-in-human robot-assisted subretinal drug delivery under local anesthesia. *Am. J. Ophthalmol.* **237**, 104–113 (2021).
19. M. Zhou, K. Huang, A. Eslami, H. Roodaki, D. Zapp, M. Maier, C. P. Lohmann, A. Knoll, M. A. Nasseri, "Precision needle tip localization using optical coherence tomography images for subretinal injection," in *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2018), pp. 4033–4040.
20. M. Zhou, J. Wu, A. Ebrahimi, N. Patel, Y. Liu, N. Navab, P. Gehlbach, A. Knoll, M. A. Nasseri, I. Iordachita, Spotlight-based 3D instrument guidance for autonomous task in robot-assisted retinal surgery. *IEEE Robot. Autom. Lett.* **6**, 7750–7757 (2021).
21. P. Zhang, J. W. Kim, P. Gehlbach, I. Iordachita, M. Kobilarov, Autonomous needle navigation in subretinal injections via iOCT. *IEEE Robot. Autom. Lett.* **9**, 4154–4161 (2024).
22. C. He, E. Yang, N. Patel, A. Ebrahimi, M. Shahbazi, P. Gehlbach, I. Iordachita, Automatic light pipe actuating system for bimanual robot-assisted retinal surgery. *IEEE/ASME Trans. Mechatron.* **25**, 2846–2857 (2020).
23. J. W. Kim, S. Wei, P. Zhang, P. Gehlbach, J. U. Kang, I. Iordachita, M. Kobilarov, Towards autonomous retinal microsurgery using RGB-D images. *IEEE Robot. Autom. Lett.* **9**, 3807–3814 (2024).
24. J. T. Wilson, M. J. Gerber, S. W. Prince, C.-W. Chen, S. D. Schwartz, J.-P. Hubschman, T.-C. Tsao, Intraocular robotic interventional surgical system (IRISS): Mechanical design, evaluation, and master-slave manipulation. *Int. J. Med. Robotics Comput. Assist. Surg.* **14**, e1842 (2018).
25. M. J. Gerber, J.-P. Hubschman, T.-C. Tsao, Automated retinal vein cannulation on silicone phantoms using optical-coherence-tomography-guided robotic manipulations. *IEEE/ASME Trans. Mechatron.* **26**, 2758–2769 (2021).
26. B. Keller, M. Draelos, K. Zhou, R. Qian, A. N. Kuo, G. Konidaris, K. Hauser, J. A. Izatt, Optical coherence tomography-guided robotic ophthalmic microsurgery via reinforcement learning from demonstration. *IEEE Trans. Robot.* **36**, 1207–1218 (2020).
27. Y.-Q. Chen, J.-W. Tao, L.-Y. Su, L. Li, S.-X. Zhao, Y. Yang, L.-J. Shen, Cooperative robot assistant for vitreoretinal microsurgery: Development of the RVRMS and feasibility studies in an animal model. *Graefes Arch. Clin. Exp. Ophthalmol.* **255**, 1167–1171 (2017).

28. C. Yan, M. Liu, G. Shi, J. Fan, Y. Li, S. Wu, J. Hu, Design of a subretinal injection robot based on the RCM mechanism. *Micromachines* **14**, 1998 (2023).
29. P. M. Łajczak, Z. Nawrat, Sharper vision, steady hands: Can robots improve subretinal drug delivery? Systematic review. *J. Robot. Surg.* **18**, 235 (2024).
30. A. Ebrahimi, M. G. Urias, N. Patel, R. H. Taylor, P. Gehlbach, I. Iordachita, Adaptive control improves sclera force safety in robot-assisted eye surgery: A clinical study. *IEEE Trans. Biomed. Eng.* **68**, 3356–3365 (2021).
31. A. Alamdar, N. Patel, M. Urias, A. Ebrahimi, P. Gehlbach, I. Iordachita, Force and velocity based puncture detection in robot assisted retinal vein cannulation: In-vivo study. *IEEE Trans. Biomed. Eng.* **69**, 1123–1132 (2022).
32. A. Ebrahimi, F. Alambeigi, S. Sefati, N. Patel, C. He, P. Gehlbach, I. Iordachita, Stochastic force-based insertion depth and tip position estimations of flexible FBG-equipped instruments in robotic retinal surgery. *IEEE/ASME Trans. Mechatron.* **26**, 1512–1523 (2021).
33. E. Svoboda, Your robot surgeon will see you now. *Nature* **573**, S110–S111 (2019).
34. P. E. Dupont, B. J. Nelson, M. Goldfarb, B. Hannaford, A. Menciassi, M. K. O'Malley, N. Simaan, P. Valdastri, G.-Z. Yang, A decade retrospective of medical robotics research from 2010 to 2020. *Sci. Robot.* **6**, eabi8017 (2021).
35. P. Fiorini, K. Y. Goldberg, Y. Liu, R. H. Taylor, Concepts and trends in autonomy for robot-assisted surgery. *Proc. IEEE* **110**, 993–1011 (2022).
36. M. Yip, S. Salcudean, K. Goldberg, K. Althoefer, A. Menciassi, J. D. Opfermann, A. Krieger, K. Swaminathan, C. J. Walsh, H. Huang, I.-C. Lee, Artificial intelligence meets medical robotics. *Science* **381**, 141–146 (2023).
37. M. H. L. Liow, P. L. Chin, H. N. Pang, D. K.-J. Tay, S.-J. Yeo, THINK surgical TSolution-One® (Robodoc) total knee arthroplasty. *SICOT-J* **3**, 63 (2017).
38. H. Saeidi, J. D. Opfermann, M. Kam, S. Wei, S. Leonard, M. H. Hsieh, J. U. Kang, A. Krieger, Autonomous robotic laparoscopic surgery for intestinal anastomosis. *Sci. Robot.* **7**, eabj2908 (2022).
39. E. De Momi, A. Segato, Autonomous robotic surgery makes light work of anastomosis. *Sci. Robot.* **7**, eabn6522 (2022).
40. K. Ray, Autonomous robotic laparoscopic gastrointestinal surgery. *Nat. Rev. Gastroenterol. Hepatol.* **19**, 148 (2022).
41. G. Fagogenis, M. Mencattelli, Z. Machaidze, B. Rosa, K. Price, F. Wu, V. Weixler, M. Saeed, J. E. Mayer, P. E. Dupont, Autonomous robotic intracardiac catheter navigation using haptic vision. *Sci. Robot.* **4**, eaaw1977 (2019).
42. R. Kline, Autonomous piloting in a beating heart. *Science* **364**, 447.20-449 (2019).
43. A. Kuntz, M. Emerson, T. E. Ertop, I. Fried, M. Fu, J. Hoelscher, M. Rox, J. Akulian, E. A. Gillaspie, Y. Z. Lee, F. Maldonado, R. J. Webster III, R. Alterovitz, Autonomous medical needle steering in vivo. *Sci. Robot.* **8**, eadf7614 (2023).
44. J. W. B. Kim, J.-T. Chen, P. Hansen, L. X. Shi, A. Goldenberg, S. Schmidgall, P. M. Scheikl, A. Deguet, B. M. White, D. R. Tsai, R. J. Cha, J. Jopling, C. Finn, A. Krieger, SRT-H: A hierarchical framework for autonomous surgery via language-conditioned imitation learning. *Sci. Robot.* **10**, eadt5254 (2025).
45. Z.-L. Ni, X.-H. Zhou, G.-A. Wang, W.-Q. Yue, Z. Li, G.-B. Bian, Z.-G. Hou, SurgiNet: Pyramid attention aggregation and class-wise self-distillation for surgical instrument segmentation. *Med. Image Anal.* **76**, 102310 (2021).
46. G.-B. Bian, L. Zhang, H. Chen, Z. Li, P. Fu, W.-Q. Yue, Y.-W. Luo, P.-C. Ge, W.-P. Liu, Motion decoupling network for intra-operative motion estimation under occlusion. *IEEE Trans. Med. Imaging* **42**, 2924–2935 (2023).
47. X. Guo, J. Zhao, L. Sun, V. Gupta, L. Du, K. Sharma, A. Van Vleck, K. Liang, L. Cao, L. Kong, Y. Yang, Y. Huang, A. De La Zerda, G. Jin, Visualizing cortical blood perfusion after photothrombotic stroke in vivo by needle-shaped beam optical coherence tomography angiography. *PhotoniX* **5**, 7 (2024).
48. N. Ciarmatori, M. Pellegrini, F. Nasini, P. M. Talli, L. Sarti, M. Mura, The state of intraoperative OCT in vitreoretinal surgery: Recent advances and future challenges. *Tomography* **9**, 1649–1659 (2023).
49. S. Dehghani, M. Sommersperger, P. Zhang, A. Martin-Gomez, B. Busam, P. Gehlbach, N. Navab, M. A. Nasseri, I. Iordachita, "Robotic navigation autonomy for subretinal injection via intelligent real-time virtual iOCT volume slicing," in *Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2023), pp. 4724–4731.
50. M. Sommersperger, S. Dehghani, P. Matten, H. Roodaki, N. Navab, "Uncertainty-aware contextual visualization for human supervision of OCT-guided autonomous robotic subretinal injection," in *Proceedings of the 2024 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2024), pp. 1268–1275.
51. X. Li, Y. Huang, Q. Hao, Automated robot-assisted wide-field optical coherence tomography using structured light camera. *Biomed. Opt. Express* **14**, 4310–4325 (2023).
52. Y. Koyama, M. M. Marinho, M. Mitsuishi, K. Harada, Autonomous coordinated control of the light guide for positioning in vitreoretinal surgery. *IEEE Trans. Med. Robot. Bionics* **4**, 156–171 (2022).
53. N. Wang, X. Zhang, D. Stoyanov, H. Zhang, A. Stilli, Vision-and-force-based compliance control for a posterior segment ophthalmic surgical robot. *IEEE Robot. Autom. Lett.* **8**, 6875–6882 (2023).
54. N. Wang, X. Zhang, S. Bano, D. Stoyanov, H. Zhang, A. Stilli, Vision-based automatic control of a surgical robot for posterior segment ophthalmic surgery. *IEEE Trans Autom Sci Eng* **22**, 6088–6099 (2025).
55. M. G. Cereda, S. Parrulli, Y. G. M. Douven, K. Faridpooya, S. Van Romunde, G. Hüttmann, T. Eixmann, H. Schulz-Hildebrandt, G. Kronreif, M. Beelen, M. D. De Smet, Clinical evaluation of an instrument-integrated OCT-based distance sensor for robotic vitreoretinal surgery. *Ophthalmol. Sci.* **1**, 100085 (2021).
56. Z. Li, Y. Deng, G.-B. Bian, W. Yu, Z. Yu, Dynamic detection of global micro-features in multimodal retinal images. *IEEE Trans. Instrum. Meas.* **73**, 4508814 (2024).
57. I. U. Scott, P. A. Campochiaro, N. J. Newman, V. Biousse, Retinal vascular occlusions. *The Lancet* **396**, 1927–1940 (2020).
58. Y. Doi, Y. Muraoka, A. Tsujikawa, Evaluation of the efficiency and safety of a 27-gauge 20,000 cuts per minute vitreous cutter. *Clin. Ophthalmol.* **17**, 2037–2043 (2023).
59. Z. Zhang, Q. Liu, Y. Wang, Road extraction by deep residual U-Net. *IEEE Geosci. Remote Sens. Lett.* **15**, 749–753 (2018).
60. Z. Li, Y. Deng, Q. Ye, W. Yu, H. Qi, Y. Liu, Z. Yu, G.-B. Bian, "Procedure recognition by knowledge-driven segmentation in robotic-assisted vitreoretinal surgery," in *Proceedings of the 2024 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2024), pp. 10779–10785.
61. Y. Diebold, L. García-Posadas, Ex vivo applications of porcine ocular surface tissues: Advancing eye research and alternatives to animal studies. *Histol. Histopathol.* **40**, 1139–1151 (2025).
62. Y. Xiao, Y. Wang, F. Li, T. Lin, K. Huffman, S. Landeros, B. Bosse, Y. Jing, D.-U. Bartsch, S. Thorogood, W. R. Freeman, L. Cheng, Acute rabbit eye model for testing subretinal prostheses. *Transl. Vis. Sci. Technol.* **8**, 20 (2019).
63. G. Nair, M. Kim, T. Nagaoka, D. E. Olson, P. M. Thulé, M. T. Pardue, T. Q. Duong, Effects of common anesthetics on eye movement and electroretinogram. *Doc. Ophthalmol.* **122**, 163–176 (2011).

Acknowledgments

Funding: This research is supported by the National Natural Science Foundation of China (grant 62027813), the National Key Research and Development Program of China (grant 2022YFB4702900), the Beijing Municipal Natural Science Foundation (L232038), and the Excellent member of CAS Youth Innovation Promotion Association (Y2022054). **Author contributions:** G.-B.B., Y.D., and Z.L. conceived the working hypothesis, proposed the methodology, and designed the experimental setup. Q.Y., Y.Z., Y.H., and Y.X. were responsible for the design and implementation of the hardware. Y.D., Z.L., and Y.Z. implemented the proposed navigation and control algorithms. Y.D., Y.Z., Y.X., W.Y., and Z.W. conducted the experiments and extracted the results. G.-B.B., Y.H., and W.Y. provided the funding. G.-B.B., Z.L., Y.H., W.Y., and Z.Y. supervised the work. All authors contributed to the writing of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data, code, and materials availability:** All data needed to evaluate the conclusions in the paper are present in the main text or the Supplementary Materials. The datasets and code used to generate Figs. 5 to 8 are available at Zenodo: https://doi.org/10.5281/zenodo.17098003.

Submitted 26 March 2025
Accepted 9 December 2025
Published 14 January 2026
10.1126/scirobotics.adx7359

Science Robotics

Autonomous robotic intraocular surgery for targeted retinal injections

Gui-Bin Bian, Yawen Deng, Zhen Li, Qiang Ye, Yupeng Zhai, Yong Huang, Yingxiong Xie, Weihong Yu, Zhangwanyu Wei, and Zhangguo Yu

Sci. Robot. **11** (110), eadx7359. DOI: 10.1126/scirobotics.adx7359

Science Robotics (ISSN 2470-9476) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Robotics* is a registered trademark of AAAS.